

# The effectiveness of a text messaging programme (Tips by Text) to promote language development in early years: randomised trial Evaluation Protocol

**Evaluator (institution): NIESR**

**Principal investigator(s): Edoardo Masset**

Template last updated: December 2022

## Evaluation summary

<b>Project title</b>	<b>The effectiveness of a text messaging programme (Tips by Text) to promote language development in early years: randomised trial</b>
<b>Developer</b> <i>(Institution)</i>	The Behavioural Insights Team (BIT)
<b>Evaluator</b> <i>(Institution)</i>	The National Institute of Economic and Social Research (NIESR) and the Institute of Employment Studies (IES)
<b>Principal investigator(s)</b>	Edoardo Masset
<b>Protocol author(s)</b>	Edoardo Masset, Janine Boshoff, Ceri Williams, Susie Bamford, Becci Newton
<b>Trial design</b>	Two-arm individual-level randomised controlled trial
<b>Trial type</b>	Effectiveness
<b>Pupil age range and Key stage</b>	4-5, Early Years Foundation Stage (Reception)
<b>Number of schools</b> <i>(at design stage)</i>	145
<b>Number of pupils</b> <i>(at design stage)</i>	3,770
<b>Primary outcome measure and source</b>	Language Development (Language Screen)
<b>Secondary outcome measure and source</b>	Self-regulation (Head-Toes-Knees-Shoulders Revised (HTKS-R))

## Protocol version history

Version	Date	Reason for revision
1.0 [original]	08/10/2024	n.a.

## Contents

Evaluation summary.....	1
Protocol version history .....	2
Study rationale and background.....	4
Intervention .....	9
Impact evaluation design .....	14
Implementation and process evaluation (IPE) design.....	29
Cost evaluation design.....	45
Ethics and registration.....	47
Data protection.....	47
Personnel.....	49
Risks .....	50
Timeline .....	52
References .....	60
Appendix 1: Tips by Text data collection plan.....	58
Appendix 2: Gantt chart of baseline data collection.....	59
Appendix 3: Tips by Text data communication plan.....	60

## Tables

Table 1: Trial design .....	15
Table 2: Ideal subdivision of schools recruited across regions.....	16
Table 3: Sample size calculations .....	20
Table 4: Broad outcomes of parents' surveys .....	35
Table 5: IPE methods overview .....	43
Table 6: Risks .....	52
Table 7 Timeline .....	54

## Figures

Figure 1: Example of a "Tip by Text" .....	9
Figure 2: Logic model of Tips by Text .....	13
Figure 3: MDES by number of schools .....	22

## Study rationale and background

There is ample evidence that early years interventions produce long-term benefits on multiple cognitive and non-cognitive skills. The promotion of cognitive and socio-emotional skills of young children, particularly among disadvantaged groups, has the potential to improve their educational outcomes, reduce crime, and ultimately increase workers productivity (Heckman and Masterov, 2007).

In the UK, early childhood policies over the last 20 years have primarily focused on service delivery and financial support (Cattan *et al.*, 2022). The rationale behind these policies is that child development is prevented by a lack of services and by a lack of parents' resources to access existing services. Over the last 20 years, governments have invested in universal family programmes providing services to parents through community centres (namely Sure Start), and in a series of tax credit schemes to support parents financially (such as the Tax-Free Childcare entitlement scheme). However, the child development literature has emphasised that public childcare services and financial support are not necessarily the most effective means to promote child development. Several studies have shown that parenting interventions supporting home learning of children at a very young age can be very effective (Attanasio, Cattan and Meghir, 2022).

A large body of evidence summarised in a series of meta-analyses shows that parents can support early literacy development through simple activities such as making conversation with children, reading books together, talking about letters and prints, reciting songs or rhymes, and through conversations during everyday activities (Mol *et al.*, 2008; Reese, Sparks and Leyva, 2010; Jaynes, 2012). Some programmes improve language development through home visits by coaches. Other programmes provide collective coaching in paediatric rooms or train parents in schools. Home visits programmes are resource-intensive and tend to be expensive, while group training is less expensive but relies on parents' attendance and does not normally reach all of the target population.

In recent years, researchers and educators have explored technology-based interventions that are relatively inexpensive and that can potentially reach a large proportion of the population. These interventions have the same goals as home visits and group training interventions but employ new technologies such as online tools, mobile applications, and text messaging. In the last ten years several interventions have used text messages to remind parents of specific actions (for example registering for school activities) or to suggest specific parenting practices. For example, Mayer *et al.* (2020) used a randomised experiment to assess the impact of some behavioural tools (including text messages) to increase the time that parents spend reading with their children. They found that parents of 3-5 years old enrolled in a subsidised pre-school programme in Chicago increased reading time by 1SD. Hurwitz *et al.* (2015) randomly allocated parents of disadvantaged children participating in the Head Start program in the American Mid-West to receive either text messages promoting parent-child interaction or to a control group. They found that parents in the intervention group changed parenting practices in ways that were more conducive to child development. Smythe-Leistico *et al.* (2018) conducted a synthetic control analysis of an intervention promoting kindergarten attendance in one elementary state school in Pittsburgh, US. To promote attendance, parents were encouraged and given feedback through text messaging. The authors found that, as a result

of the programme, absenteeism in the school decreased considerably in comparison to other schools.

A study of a text messaging intervention by York et al. (2019) found significant improvements in children's language skills and was replicated and expanded in new directions by other researchers. York et al. (2019) randomly assigned pre-kindergarten families of an educational district of San Francisco, to receive 3 weekly messages over 8 months, promoting home learning, or to a control group. Text messages belonged to one of three categories: 1) "fact" texts designed to inform and motivate parents, 2) "tip" texts suggesting specific activities building on standard family routines, and 3) "growth" texts providing parents with encouragement and reinforcement. The study found that at the end of the intervention children in the intervention group scored significantly better on a standard literacy test, and that improvements were particularly large for children that had scored more poorly at the baseline before the start of the intervention.

Cabell et al. (2019) replicated the intervention studied by York et al. in a suburban setting in a mid-Atlantic state using a relatively smaller sample of parents and children, and found no positive effects on any of the literacy measures employed. Cortes et al. (2021) extended the study by York et al. by testing whether the number of messages could produce a smaller or larger impact. They randomly allocated parents of 4-year old pre-schoolers in the Dallas independent district, to three arms receiving either 1 message per week, 3 messages per week or 5 messages per week. They found no additional gains in literacy outcomes from the higher or lower dosage of messages, and that fewer or more messages were producing poorer results for some subgroups of the sample. Doss et al. (2019) extended the study by York et al. by comparing the original intervention to another text messaging intervention that provided personalised messages to parents including information on their skills and grades. They conducted the experiment among the same population of the original York et al. study, now in kindergarten, and found that while literacy did not improve in the arm receiving the original intervention, there were significant effects of the personalised intervention. Finally, Doss et al. (2022) conducted a three-arm trial among parents of pre-schoolers in three California school districts, which compared a) text messaging promoting parents' behaviours to improve maths skills, b) text messaging promoting maths, literacy, and socio-emotional learning at the same time, and c) a control group. They found that the pure maths intervention had no impact, while the combined intervention produced an impact on children's maths test scores but only for specific subgroups (girls and girls within a specific range of the outcome distribution).

The study by York et al. (2019) showed that text messaging interventions can be an effective way to improve children's language skills. The extensions of the study conducted by other researchers in other settings helped to refine the intervention in order to maximise its impact. The Behavioural Insights Team (BIT) adapted the curriculum of the original text messaging intervention in York et al. to the British context. They modified the content of the messages to be aligned with the Early Years Foundation stage profile, and they successfully tested it in 109 schools in the North East of England.

An efficacy evaluation of the intervention designed by BIT was conducted by a team comprising NIESR and IES over the period from 2019 to 2021 and with funding from EEF. The study was designed as a parent-level randomised trial including 109 schools and 2,647 children located in the North-East of England. The primary outcome of interest was the YARC

literacy test, and secondary outcomes included the Total Difficulty score and the Strengths and Difficulties Questionnaire. The evaluation found an impact of the intervention of 0.05SD on the YARC literacy test and between -0.02 and 0.08SD on Strength and Difficulty outcomes. None of the results were statistically significant. The impact on the literacy score was 0.09SD for beneficiaries of FSM and 0.04SD for the half of the sample that performed more poorly at the baseline, but again both results were not statistically significant.

Unfortunately, the trial was heavily disrupted by COVID-19. Schools were closed between March and July 2020, and again between January and March 2021. The closures forced the delivery team to extend the duration of the intervention, while children's assessments were postponed and delayed. In the end, only 753 children were tested at the endline, implying an attrition rate of 72% mostly resulting from the disruption produced by COVID-19. This impacted the ability of the study to detect a statistically significant effect of the intervention, while the intervention was not carried out in the desired way.

The IPE study found that the intervention was well accepted by parents and correctly implemented by the delivery team. Compliance was high, as parents received on average 91% of the messages sent. Few parents withdrew from the intervention and 38% reported using the message tips regularly while 60% reported some use. Parents reported that they found the messages easy to understand and at the right level for their children's age and abilities. Most parents thought the intervention improved their children's language skills and the relationship with their children.

Surveys with teachers showed that they were blind to project assignment and largely unaware of the contents of the text messages or of the intervention taking place. Interviews with teachers revealed that in the business-as-usual case, teachers already promote home learning activities for parents through meetings, web applications, and text messages, but these meetings are few, the information content is limited and the use of applications such as *Tapestry* or *ClassDojo* is unclear. In other words, it seems that outside of Tips by Text, very little support for home learning was offered to parents.

One issue that emerged during the efficacy trial was the risk of contamination of the control group (Stokes *et al.*, 2022). Parents participating in the intervention reported sharing text messages with other parents in the following proportions: 22% (150 parents) sharing with parents of children in other schools; 11% (72 parents) sharing with parents of children in same school but different year; 21% (141 parents) sharing with parents of children in same school and year; and 47% (318 parents) not sharing. Worryingly, up to a fifth of parents shared text messages with parents of children attending the same school and year, and therefore potentially with parents of children included in the control group. During their interviews with the evaluation team some parents suggested that COVID-19 may have led them to increase message sharing, whereas other parents suggested that COVID-19 may have led them to decrease sharing, so that the overall impact of COVID-19 on message sharing is unclear. As for the intensity of contamination, some parents reported sharing all messages, while others reported sharing some of them. If parents of children in the control group were receiving text messages, the impact of the intervention could have been underestimated.

The outcomes of the efficacy trial produced some lessons that were incorporated into the design of the current study. The report of the Tips by Text efficacy trial reached five key conclusions, and some additional lessons learned.

- First, the attrition rate was extremely high (70%) mostly because of the disruption produced by COVID-19. The researchers concluded that the original research questions could only be answered in a time less acutely affected by COVID-19 or other shocks.
- Second, the efficacy trial found a small non-statistically significant impact of the intervention on literacy as measured by the YARC assessment. This led the researchers to review the outcome measures and to select a primary outcome indicator (LanguageScreen), which is less sensitive to floor effects, and to include other intermediate outcome indicators that allow assessing the impact of the intervention along the theory of change such as, for example, changes in parents' engagement with schools and changes in the home learning environment. The researchers also concluded that the detection of small impacts such as those produced by the efficacy trial would require a large sample size.
- Third, the study found some positive, although not statistically significant, effects of the intervention on FSM eligible children. This is aligned with other studies of text messaging interventions, which often find larger impacts among disadvantaged children. The researchers decided to expand the sub-group analysis of the study by including an analysis by EAL status and an analysis based on pre-intervention test scores.
- Fourth, the IPE concluded that the intervention was well accepted by parents and children and that compliance was high. Parents agreed that messages were delivered at the right level, time, and with the adequate intensity. As a result, no changes in project delivery have been considered and the acceptability of the intervention is not a key question of current IPE.
- Fifth, the IPE concluded, as discussed above, that significant sharing of text messages occurred between parents resulting in potential contamination of the control group. The researchers responded to the risk in the current study by a) increasing the size of the sample to capture an impact of the intervention biased downwards by contamination, and b) including questions in the parent surveys to capture the extent of sharing and allowing a 2-sided compliance analysis of the impact of the intervention that estimates the impact of the intervention at the same time accounting for contamination.

In light of these conclusions, we identified three key learning points. First, the original intervention design appears to be both feasible and acceptable. Parents reported that the number of messages and their contents were appropriate. Compliance was high. No major change to project delivery was required. Second, the expected effect size of the intervention is small and requires a large sample size, particularly for estimating the impact of the intervention on sub-groups of interest. In the efficacy trial the intervention produced a population-level impact of 0.05SD, which using the standard Cohen's rule of thumb (<0.20SD) qualifies as a small effect. The need for a large sample size informed the power calculations of the current protocol and it is discussed in the section on Impact Evaluation Design. Third, the potential risk of contamination of the control group has a series of implications for the delivery of the intervention, the sample size, and the analysis of the results that are discussed throughout the report.

The evaluation is mixed-method and will comprise a randomised controlled trial, which will enable examination of the quantitative impact of the programme on children's language skills, and an impact and process evaluation (IPE), to understand the extent to which the programme is implemented as intended (fidelity), and to identify the mechanisms through which the intervention operates as spelled out in the theory of change (impact mechanisms).

The study is an effectiveness trial with assignment at the individual (family) level. A total of 145 schools (spanning four large geographic regions representing approximately 50% of England's population) are included in the study. The average reception class size in the academic year 2022/23 was 26.6.<sup>1</sup> For the purpose of the trial, the study will include one class in each school, and all children in that class, which will deliver a sample of approximately 3770 children. Half the pupils, in each selected class within each school, will be assigned to the intervention group. The intervention is delivered to parents, and school staff and the data collection team will be blind to the assignment. The evaluation will assess the impact of the intervention on language development (using the Language Screen assessment tool) and on self-regulation (using the Head-Knees-Toes and Shoulder assessment tool). In addition, changes in parents' behaviour in the home learning environment will be measured to explain the effectiveness of the intervention. Mechanisms of impact will be further investigated by an Impact and Process Evaluation, which will also explore issues of fidelity, adaptation, project differentiation, and contamination, by conducting surveys and qualitative interviews with parents and surveys with school staff. The intervention will run for nine months starting in November 2024 and ending in July 2025, surveys of parents and children's assessments will be conducted at baseline and endline.

It will also aim at employing secondary administrative data from the National Pupil Database (NDP) to detect impacts of the intervention on pupil sub-groups and in relation to outcomes not directly measured by our study. Text messaging interventions such as Tips by Text are low-cost in comparison to other parenting interventions such as for example home visits or collective coaching in dedicated settings. The evaluation will also calculate the cost of the programme.

A randomised trial is the ideal method to assess the effectiveness of the intervention in an unbiased way. Other evaluation methods relying on the analysis of groups self-selected in the intervention would be potentially biased because they are unable to separate the effects of the intervention from the effects of characteristics determining the participation in the intervention. The evaluation is designed as an individual-level randomised trial rather than a cluster-level trial for pragmatic reasons. In principle, a cluster-level level trial would be more appropriate to rule out the risk of contamination of the control group. However, a cluster-level trial is not feasible as the text messaging interventions produce small or very small effects which require large samples. A cluster randomised trial would require a sample size which is nearly twice the size of the sample required by an individual-level trial, entailing a prohibitive cost.

---

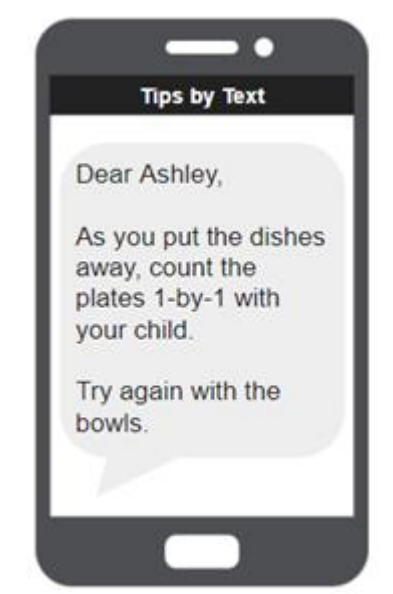
<sup>1</sup> <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics>

## Intervention

### Name of the intervention: Tips by Text

Tips by Text is a nine-month curriculum adapted in line with the Early Years Foundation Stage Profile (EYFSP). Parents will receive three text messages weekly, sent by BIT, which contain tips for specific and easy to implement activities designed to improve language, literacy, numeracy and socio-emotional skills. Each week messages will centre on a particular topic like letter recognition, counting or managing emotions. Figure 1 shows an example of the text messages sent to parents.

Figure 1 Example of a “Tip by Text”



**Why:** early years parenting interventions have the potential to substantially improve children’s cognitive abilities and socio-emotional skills. Providing parents with simple instructions on parenting practices through text messages has the potential to reach a large population and can be carried out at low-cost. Recent research in the United States has found that this type of intervention is effective in improving academic achievements and socioemotional skills, particularly among disadvantaged children and communities.

**Who (recipients):** half parents of pupils in Reception year (aged 4 to 5) in recruited school. It is only one parent per child (the school’s primary parent contact) that will receive the intervention.

**What (materials):** parents allocated to the intervention receive three text messages (delivered by phone) per week for nine months (including during school holidays) over the period from November 2024 to July 2025. There are three types of messages: 1) “FACT” texts inform and motivate parents by highlighting the importance of a particular skill or set of skills; 2) “TIP” texts minimise the cognitive, emotional, and time burden of parents by recommending short, simple, and highly specific activities to do with their children that build on existing family routines; 3) “GROWTH” texts provide encouragement and reinforcement and extend the content of the “TIP” texts.

**Who (provider):** BIT developed the text-messages and will set up the delivery of the messages. They will also respond to any messages received from parents or from the schools and will alert teachers of any safeguarding concerns that might arise.

**How (format):** text messages are sent on Tuesday (3:45PM), Thursday (3:45PM), and Saturday (10:30AM) in a standard text format using the texting platform FireText . If delivery fails, another attempt is made the following business day. If four messages in a row are undelivered (four messages and four resends) the delivery team will contact the school to check that the parent phone number is correct.

**Where (location):** schools are selected from four geographic regions of England: North West, Yorkshire and The Humber, London, and the South East.

**When and how much (dosage):** three messages per week will be sent over a period of nine months running from November 2024 to July 2025 (including during school holidays).

**Adaptation:** Parents can adapt the activities suggested by the text messages to suit them and their circumstances. The text messages and the curriculum are final and not expected to change and will not be adapted to parents' feedback or response behaviours. However, the delivery team will be able to alter the content of the messages if exceptional circumstances would require. For example, during the efficacy trial conducted during the COVID-19 pandemic, the delivery team was able to quickly adjust the content of the messages to government restrictions on socialising and on time spent outside homes.

### **Theory of change**

The theory of change of the intervention is illustrated in the logical framework in Figure 2. The intervention is designed to address a series of cognitive barriers that lead parents to underinvest in the academic development of their children. Four key cognitive barriers are identified: insufficient information, task complexity, limited attention, and time-inconsistent preferences (York, Loeb and Doss, 2019). First, parents do not have comprehensive information about their children's skills, how skills can be improved, and what skills matter for the cognitive and socioemotional development of the child. Second, parenting is an extremely complex task, and parents may resort to simple heuristics or habits when making important choices about the future of their children. There is a risk that parents make suboptimal choices. Third, parents suffer from limited attention because of the multiple demands on their time and the difficulty of balancing work and family. This limited attention poses a further risk on the ability of making optimal choices. Fourth, individuals tend to discount the future more heavily than the present and therefore tend to make choices that are biased towards more immediate rewards rather than distant ones. Long term educational choices, which offer large future rewards at a present cost, are negatively affected by parents' time preferences.

It is important to stress that the four cognitive biases described disproportionately impact families that are at a disadvantage or who live in disadvantaged communities. Poverty negatively affects parents' education and their access to information. Poverty also reduces attention and limits the ability of addressing complex tasks because meeting daily challenges such as paying rent and providing food constrains the ability to address more difficult choice tasks (Haushofer and Fehr, 2014). Finally, poor people tend to discount the future more heavily because they are pressed to find solutions to more immediate problems related to basic

survival. Studies of behavioural change in education have often found that interventions acting on parents' behaviour are more effective for disadvantaged or minority pupils (Damgaard and Nielsen, 2018).

Reviews of behavioural interventions in education have found that they tend to benefit mostly those affected by the specific barrier addressed by the intervention (Damgaard and Nielsen, 2018). Tips by Text targets four cognitive barriers. Every Tuesday, parents receive a FACT message. The FACT message highlights the importance of a particular skill or set of skills and is designed to address parents' information deficit. On Thursday they receive a TIP message. The TIP message suggests simple and very specific activities for parents to carry out with their children that build on existing family routines and is designed to address parents' attention gap and the complexity of parenting. Finally, on Saturday parents receive a GROWTH message. GROWTH messages provide encouragement and motivation to parents and are designed to address time inconsistencies and short planning horizons.

One key assumption of the theory of change behind Tips by Text is that early years parenting interventions are effective in improving children's cognitive and socioemotional skills, but that they are demanding in terms of parents and staff time and unable to reach the entire target population. The use of text messaging by Tips by Text is meant to be a low-cost way of providing support to parenting that can reach a large fraction of the target population. It should be emphasised that text messaging interventions are not expected to produce large impacts, but are designed to operate at relatively low cost per pupil. This highlights the need to collect accurate cost data for the intervention in order to allow a future cost-effectiveness analysis.

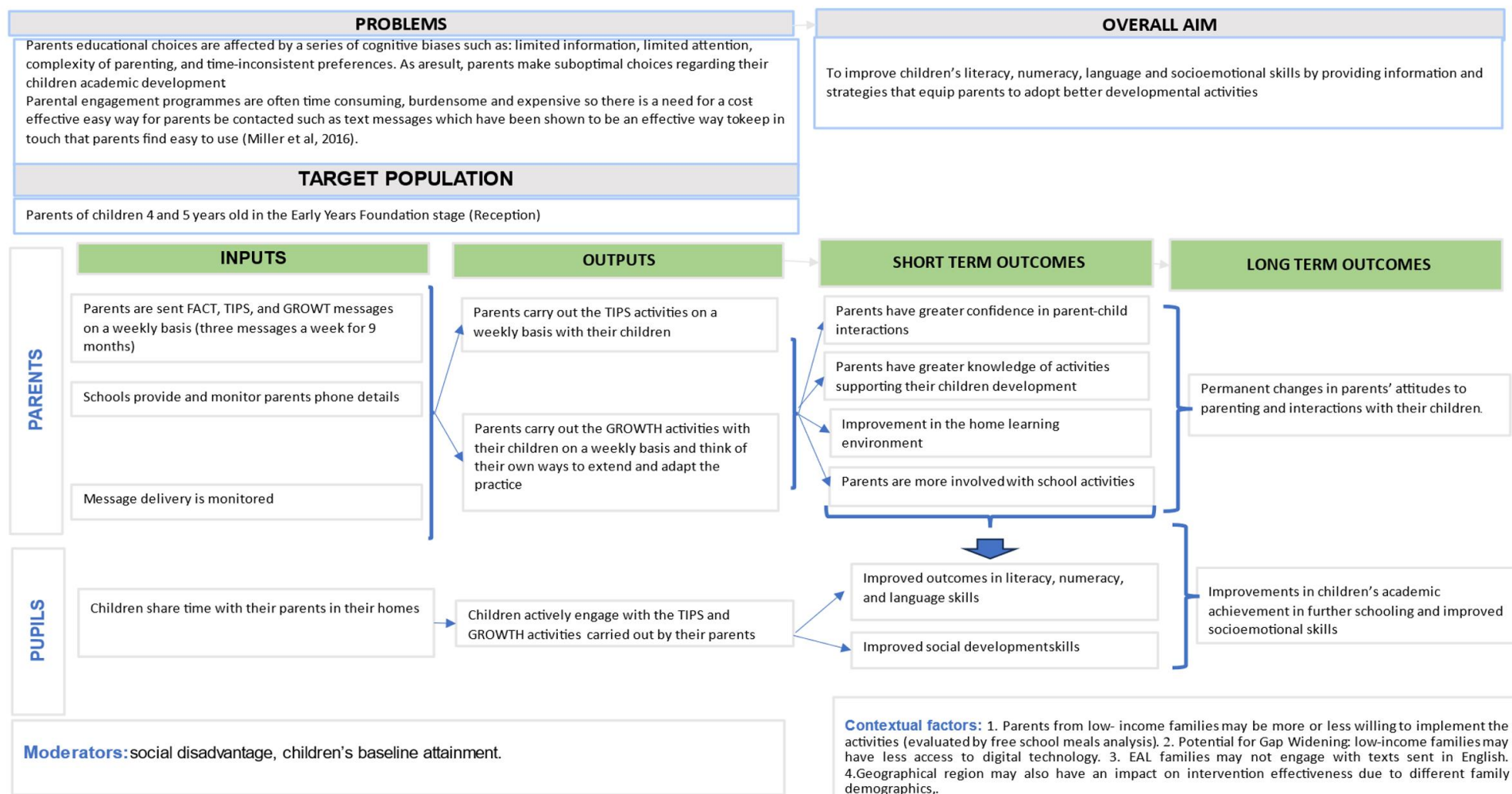
The inputs required by the intervention are few and the implementation is relatively straightforward. Tips by Text mostly consists of delivering text messages via phones. Limited follow up is needed, either with the parents or with the school, and it mostly consists of reminders and of updating contact details. The educational project activities are effectively carried out by the parents involved. The success of the intervention lies in its ability to motivate parents to actively engage with their children's learning. Parents follow the "tips" provided by the intervention, and are encouraged to develop their own methods to stimulate their children's learning. The project hopes in this way to change overall parents' attitudes to children learning. It is expected that parents will become more interested in their children's learning and more involved with school activities. A change in parental attitudes towards children learning and a higher involvement with school activities are therefore intermediate impacts of the intervention. Even if this isn't the case, the intervention could still positively impact outcomes, providing parents simply engage with and act on the messages themselves.

Several factors may affect the effectiveness of the intervention either positively or negatively. First, text reminders may backfire (Damgaard and Nielsen, 2018). No such negative effects have been observed in previous studies of READY 4K conducted in the US, but it cannot be excluded that some parents may be negatively affected by the intervention. Second, parents need to have the time to put in place the suggestions provided in the text messages. The activities recommended by the messages require minimum effort and time on the part of parents, but some parents may be unable to follow the instructions provided. Third, the messages may be redundant among well-educated parents, who might be conducting many of the suggested tips already.

The intervention is universal and directed to all parents of children attending a specific setting. However, as mentioned, behavioural education interventions tend to be more effective among disadvantaged groups and when they address specific cognitive biases that are more common among people at a socioeconomic disadvantage. We should expect Tip by Text to have a larger impact among children of disadvantaged parents. This points to the need to analyse the impact of the intervention for particular sub-groups and to power the study accordingly.

Finally, although the intervention is universal, in order to assess its effectiveness through a randomised trial we have to exclude some parents from the intervention. This poses a risk of contamination of the control group. There is a possibility, highlighted in the report of the Tips by Text evaluation (Stokes *et al.*, 2022), that parents in the intervention group share the content of the messages with parents of the control group. Such sharing would not occur in a real setting where all parents receive the text messages. This in turn implies that the effect of the intervention could be underestimated as the control group could be, to some extent, positively affected by the intervention.

Figure 2: Logic model of Tips by Text



# Impact evaluation design

## Research questions

The **primary research question** that this impact evaluation is designed to address is:

RQ1. What is the difference in language development measured by the LanguageScreen assessment among 4-5 years old children in Reception Year exposed to Tips by Text, compared to those 4-5 years old children in the control group?

The **secondary research questions** this evaluation is intended to answer are:

RQ2. What is the difference in the different constitutive components of language development as measured by the LanguageScreen assessment among 4-5 years old children in Reception Year exposed to Tips by Text, compared to those 4-5 years old children in the control group, when measuring:

- a) expressive vocabulary
- b) listening comprehension
- c) receptive vocabulary
- d) sentence repetition

RQ3. What is the difference in self-regulation, as measured by the HKTS-R assessment, among 4-5 years old children in Reception Year exposed to Tips by Text, compared to those 4-5 years old children in the control group?

RQ4. What is the impact of Tips by Text on the outcomes above for a sub-group of socially disadvantaged children (FSM/EYPP) and for children with EAL?

The final research question concerns the intermediate impact of the intervention on parents' attitudes and behaviours:

RQ5 What is the impact of Tips by Text on parents' attitudes and behaviours towards home learning, and on their involvement with school activities?

## Design

This is an effectiveness trial that will examine the impact of the Tips by Text intervention under circumstances that more closely match "real-world" conditions. The goal of the trial is to determine how the intervention works if implemented under actual conditions. The trial is conducted in four regions that together represent more than half of the English population (52%), and the large number of schools within regions will account for the heterogeneity in the population. The results will be sufficiently representative of the implementation of a new similar intervention in England.

The trial is designed as a two-arm, individual-randomised trial, where the parents (and their child) are the units of randomisation. Tips by Text will be delivered to the parents of children in reception year (aged 4-5 years old). One arm of the study will receive the intervention, while the other arm will not. At the end of the programme, all schools participating in the trial will receive the text message programme free of charge.

Table 1: Trial design

<b>Trial design, including number of arms</b>		Two-arm, individual-level randomised trial
<b>Unit of randomisation</b>		Parents
<b>Stratification variables</b> (if applicable)		Schools (Reception class ages 4-5)
<b>Primary outcome</b>	<b>Variable</b>	Language development
	<b>Measure</b> (instrument, scale, source)	Expressive vocabulary, listening comprehension, receptive vocabulary, and sentence repetition as recorded by the <b>LanguageScreen</b> assessment
<b>Secondary outcome(s)</b>	<b>Variable(s)</b>	Self-regulation
	<b>Measure(s)</b> (instrument, scale, source)	<b>Heads-Toes-Knees-Shoulders Revised (HTKS-R)</b>
<b>Baseline for primary outcome</b>	<b>Variable</b>	Language development
	<b>Measure</b> (instrument, scale, source)	Expressive vocabulary, listening comprehension, receptive vocabulary, and sentence repetition as recorded by the <b>LanguageScreen</b> assessment
<b>Baseline for secondary outcome</b>	<b>Variable</b>	Self-regulation
	<b>Measure</b> (instrument, scale, source)	<b>Heads-Toes-Knees-Shoulders Revised (HTKS-R)</b>

The business-as-usual case in the control group will present considerable regional and setting heterogeneity. Home learning and parents' involvement in child learning is promoted in schools in various ways and different intensity. The efficacy trial report found that teachers in the schools surveyed by the study already promote home learning activities for parents through meetings, web applications, and text messages. However, these meetings are few, the content of the instructions to parents is limited, and the use of applications such as Tapestry of ClassDojo is limited. It is unlikely that other major interventions promoting child development will be in operation at the time of the study, and in any case the large sample and its wide geographic distribution prevents the potential that other interventions could have a major impact on the estimated impact of the intervention.

The trial aims to establish the impact of the Tips by Text programme on young children's language skills. The primary outcome is language development measured by the LanguageScreen assessment. Secondary outcomes are the four constitutive components of language development measured by LanguageScreen separately considered: and self-regulation. These measures were chosen in collaboration with the delivery team and reflect a consideration for the young age of the study participants, as well as practical factors related to minimising the burden of conducting assessments with the participants. These outcome measures are described in further detail later in this protocol.

While the intervention will be delivered to half parents of pupils in Reception year in all the recruited schools, for the purpose of the trial, only a sub-sample of pupils will be included in the sample and will be therefore be administered an assessment at baseline and endline. The section on sample size explains in detail how the sub-sample of pupils to be part of the trail is selected.

## Participant selection

The trial will be conducted in four geographic regions of England (North-West, Yorkshire and The Humber, London, and the South East). In principle, to ensure the study yields results that are representative of the population, the proportion of schools recruited in each region should align with the proportion of that region’s population in the total population involved in the study. Recent UK census data indicates that the four regions considered in our study represent 52% of the total English population. Table 2 indicates the percentage of the study population in each region and the corresponding number of schools that would be required from each region to achieve representative results.

Table 2 Ideal subdivision of schools recruited across regions

Region	Proportion of study population	Schools required to obtained representativeness
London	0.31	45
North-West	0.26	38
South East	0.23	33
Yorkshire and the Humber	0.20	29

Assuming a total of 145 schools for the study (as will be discussed in the section on statistical power), Table 2 indicates the number of schools needed in each region to obtain a regional representative sample. For pragmatic reasons the study did not aim to obtain a representative sample, but aimed at being representative as much as possible. The figures in Table 2 are an ideal subdivision of schools across regions, and it is recognised that for pragmatic reason related to school recruitment an exact match is impossible to achieve. We therefore agreed to aim to recruit an approximate number of 30-40 schools per region, with a minimum of 22 school per region, to make the sample reasonably representative at the geographic level.

Schools in these regions will receive an information sheet explaining the Tips by Text programme, the timeline over which the programme will be delivered, and the data that will be collected as part of the evaluation. Interested schools will email the delivery team directly and will be added to a list of schools that will be considered for inclusion in the trial.

The study will recruit up to 145 schools. Additional schools will be included in a reserve list to allow for replacements in case some of the 145 initially recruited schools were to withdraw from the study.

The delivery team applied as an eligibility criteria during recruitment that schools were not already taking part in any similar EEF studies in 2024/25.

The trial aims to recruit 145 schools with a minimum average of 26 students pupils per Reception year group. Half of the children in a reception-year class will be randomly assigned to receive the programme (the treatment group), while the remainder will not receive the programme (the control group).

Recruitment will be led by the Behavioural Insights Team (BIT), with support from the evaluation team. All recruitment documents are developed as a collaborative effort between the delivery team, the evaluation team and the EEF. An initial information sheet provides schools with information about the project and provides contact details for the delivery team should the school wish to be considered for inclusion in the trial. Following discussion with the delivery team, if a school is interested in proceeding, the school will be asked to sign a memorandum of understanding which sets out what participation in the project entails and the roles of all parties.

Schools will be asked to distribute information sheets to parents of eligible children along with a privacy notice. The information sheet provides parents with information on how they can withdraw their child from participating in the research activities, as well as how to withdraw their and their child's information from the evaluation activities.

## Outcome measures

### **Baseline measures**

The study will employ two assessments: the LanguageScreen test and the Heads-Toes-Knees-Shoulders Revised (HTKS-R) test. Following the causal chain presented in the theory of change, the first assessment measures the impact of the intervention on literacy and language development and the second assessment measures the impact of the intervention on social development skills as proxied by self-regulation. Both assessments will be carried out at baseline and at follow-up.

There are two reasons for administering the same tests before and after the intervention. First, pre- and post-intervention scores for the same children increase the statistical power of the study thanks to the correlation between pre- and post-interventions scores. Second, baseline values of the tests will allow the researchers to check how well balanced the observations are across the control and intervention groups. Previous experience with similar trials has shown that even large samples may deliver unbalanced groups. Information on differences in baseline test scores is needed at the analysis stage to interpret the results correctly. All the assessments will be administered by a trained team of assessors from Qa Research. An education specialist from NIESR will train Qa staff, which in turn will organise in-person training for the team of assessors. The scores of LanguageScreen are automatically generated by a specialised software and stored in a data management system that will be accessible to NIESR and Qa only. Similarly, the scores of the HTKS-R assessment will be generated using a specialised software and will be accessible to the Qa team only. They will be shared with NIESR using a secure Onedrive link within one week from completing all the assessments to allow for data quality checks.

LanguageScreen is a short and reliable automated test, which requires between 5-10 minutes to complete. It was found to have excellent psychometric properties, including high reliability, good fit to the Rasch model, and minimal differential item functioning across key student

groups (Hulme, 2023). The HKTS-R test takes about 5-8 minutes to administer, and the two assessments together take between a minimum of 10 minutes and a maximum of 20 minutes to administer. The test has been shown to have a strong construct and predictive validity (McClelland *et al.*, 2021). A meta-analysis of 69 studies employing this test found that it has positive associations with children's overall academic performance and that it was more strongly associated with children's mathematics performance relative to their performance on language and literacy measures, and that these associations remained stable across different participant and measurement factors (Kenny *et al.*, 2023).

At the point of pre-assessment, the treatment/control status of the parents of each child will not be known as randomisation will take place after the completion of pre-tests. At post-test, the assessor team will also not be aware of the treatment or control status of the parents of each child. The resulting scores will be passed to the evaluation team for analysis.

### **Primary outcome**

The LanguageScreen assessment is the primary outcome. LanguageScreen is a comprehensive assessment of children's oral language skills and tests their competency across four fundamental language skills: expressive vocabulary, listening comprehension, receptive vocabulary, and sentence repetition.

The test is easy to use and can be administered by professionals and non-professionals using a tablet. The assessment takes between 10 and 15 minutes to complete and is suitable to test language skills of children aged 3.5 to 8 years. LanguageScreen is made up of four tests which involve looking at what is shown on the screen or listening to what is said aloud by the app or the adult conducting the assessment:

- Expressive vocabulary: This test involves the child looking at 24 pictures on the app and identifying the image shown. If the pupil gets 8 answers wrong in a row, the test will end early.
- Listening comprehension: This test assesses a child's ability to follow a simple story and remember the details. The test involves the child answering 16 questions about 3 short stories.
- Receptive vocabulary: The test involves 23 sets of four pictures where the child must identify the picture that matches the word spoken out loud.
- Sentence repetition: The test involves 14 sentences read out loud which the child must then repeat verbatim; and allows us to test the child's understanding of sentence structure and meaning.

The scores are presented as both a table and a graph which provides information on the raw score (i.e. the number of correct answers) for each test, as well as their standard scores<sup>2</sup> that have been adjusted for their age based on a database of over half a million assessments. Standard scores are available for each sub-test, but the child's total summary score of LanguageScreen provides a holistic view of their language ability and is the primary outcome

---

<sup>2</sup> Standard scores range from 65 to 135, with 100 being the average score.

of the study. A percentile score is also available for each child tested, indicating their relative position in the distribution of language skills for their age.

The LanguageScreen was selected as the primary outcome measure for several reasons. First, it represents a comprehensive measure of four fundamental language skills, as well as a child's overall oral language skills. Good oral language skills are an important foundation for the development of literacy and numeracy skills in Key Stage 1 (KS1) and beyond. Second, LanguageScreen is proven to be sensitive to detecting changes in language skills of the target age group of our population (age 4 to 5). Third, the assessment has been validated for a UK population, and has a database of over a half million of completed assessments against which to benchmark (Hulme *et al.*, 2023).

The LanguageScreen is a commercial product, and thus cannot be attached in the appendix.

### **Secondary outcomes**

There is significant evidence pointing to a bidirectional relationship between language skills and executive function (Bishop, Nation and Patterson, 2014; Gooch *et al.*, 2016; Kaushanskaya *et al.*, 2017). Since Tips by Text aims to improve oral language skills through encouraging communication using language, both speaker and listener exhibit executive function by regulating their thoughts and actions according to their goals (Shokrkon and Nicoladis, 2022). To determine whether the programme has an impact on behavioural self-regulation, we propose to conduct the Heads-Toes-Knees-Shoulders Revised (HTKS-R) assessment at both baseline and endline.

HTKS-R is an assessment of children's behavioural self-regulation that requires multiple executive function components and has strong reliability and validity, as well as having predictive associations with academic and classroom outcomes (McClelland *et al.*, 2021). The HTKS-R is an extension of the original HTKS assessment that was developed to test self-regulation in children aged 4 to 8 years old. The revised version starts with a verbal section before continuing to the standard 3-part assessment in the HTKS. While the HTKS-R was developed to assess younger children (aged 3 to 8 years old), it has also been shown to reduce floor effects in older children who exhibit socio-demographic risk factors (single-parent household, parents with low educational attainment, or an ethnic minority status) or dual-language learners (Gonzales *et al.*, 2021; McClelland *et al.*, 2021). Given our interest in the impact of the programme on children from disadvantaged backgrounds, the evaluation team opted to use the HTKS-R to assess self-regulation.

The HTKS-R has been developed for research purposes only, and a condition of gaining access to the task is that researchers will not share it with colleagues outside of the evaluation team. Therefore, the HTKS-R cannot be attached in the appendix.

## Sample size

Table 3: Sample size calculations

		Overall	EYYP/FSM <sup>1</sup>	EYYP/FSM <sup>2</sup>
<b>Minimum Detectable Effect Size (MDES)</b>		0.08	0.13	0.16
<b>Pre-test/ post-test correlations</b>	level 1 (pupil)	0.40	0.40	0.40
	level 2 (class)	-	-	-
	level 3 (school)	0.10	0.10	0.10
<b>Intracluster correlations (ICCs)</b>	level 2 (class)	-	-	-
	level 3 (school)	-	-	-
<b>Alpha</b>		0.05	0.05	0.05
<b>Power</b>		0.80	0.8	0.8
<b>One-sided or two-sided?</b>		Two-sided	Two-sided	Two-sided
<b>Average cluster size</b>		20	7	5
<b>Number of schools</b>	Intervention	145	145	145
	Control	145	145	145
	<b>Total</b>	145	145	145
<b>Number of pupils</b>	Intervention	1,450	508	376
	Control	1,450	508	376
	<b>Total</b>	2,900	1,016	752

<sup>1</sup> Assuming 33% of FSM children as in the sample of recruited schools

<sup>2</sup> Assuming 26% of FSM children as in the average English school

Table 3 presents the result of the power calculations to determine the sample size of the study. The calculations were made using the Optimal Design software, using a model assuming an individual level randomisation, a significant correlation in the test results before and after the intervention for the same pupil, a moderate correlation between schools and pupils' scores, and a reasonable amount of variability of effect sizes between schools (a random effect model).

The chart below plots the minimum detectable effect size as we increase the number of schools surveyed. We employed the followed parameters. Following conventions, we set the statistical level of significance ( $\alpha$ ) of a two-tail test in the difference between means at 0.05, and we set the desired statistical power of the study ( $\beta$ ) at 0.80 (representing 80% of chances of finding an effect if there is indeed an effect).

In the statistical analysis we will use an ANCOVA model regressing the test scores after the intervention against the test scores before the intervention. The use of pre-intervention scores as an explanatory variable increases statistical power with the size of the correlation between the score before and after the intervention. We set the correlation between pre- and post-

intervention tests scores (the proportion of variation by level-1 covariate -  $R^2$ ) at 0.40. This is in accordance to the results of EEF research (Singh et al., 2023), which found the correlation between pre and post-test English tests between Early Years Foundation Stage and Key Stage 1 (the closest comparison to our assessments) in the range 0.43-0.63 using the NPD data, and in the range 0.54-0.88 using the data from past EEF-funded trials. We adopt the most conservative of these estimates.

In the statistical analysis, the test scores will be regressed against the pre-intervention scores and a series of dummy variables representing school-level effects. These dummy variables capture the average effects of characteristics of the school such as the school environment, the socio-economic profile of the population, or the quality of the teaching staff (fixed effects). We assume that the inclusion of these variables will explain a relatively small fraction of the overall variance in test scores of 0.10 (this is parameter B in Optimal Design – the proportion of the variance explained by the blocking variable).

Lastly, we assume that the effect of the intervention varies across schools as in a random effect model. We have no indications from other studies of what could be the value of this variability, but we can make some reasonable assumptions. We set the expected variability of the effect size across schools ( $\sigma^2$ ) at 0.01. This variability implies a standard deviation of 0.10 (the square root of 0.01) and, assuming an effect size of 0.10, a confidence interval of [-0.10,0.30] around the expected effect. This confidence intervals means that 95% of the school-level treatment effects are expected to be between -0.10 and 0.30, and that 25% of schools will be negatively affected by the intervention. A larger variability of effect sizes would lead to confidence intervals that are too wide and deliver some extremely large effects including many negative effects.

We set the number of children tested in each school at 20. The study will select one class within each recruited school and will target all children in that class. The size of a reception class in England is legally capped at a maximum of 30 pupils per class. However, some classes are formed with fewer than 30 children and the average reception class size in England in 2023/24 was 26.6. In addition, school recruitment aimed at selecting schools with a minimum of 26 pupils in Reception group as projected by attendance recorded in previous years. Therefore, we expect to carry out more than 26 children assessment per school. We then expect an attrition rate of about 20% resulting from missing either baseline or endline observations for reasons of school absence, refusal or just measurement error. We therefore use in the calculations the more conservative number of 20 children per school because although our intention is to test all children included in the study, we are aware of a relatively high risk of attrition across surveys.

Some schools are multi-class and have more than one class. In the case of multi-class schools, the evaluation team will randomly select one class for the study. This random selection will occur before the collection of pupil details data from the schools in September.

Figure 3 Minimum detectable effect size by number of schools

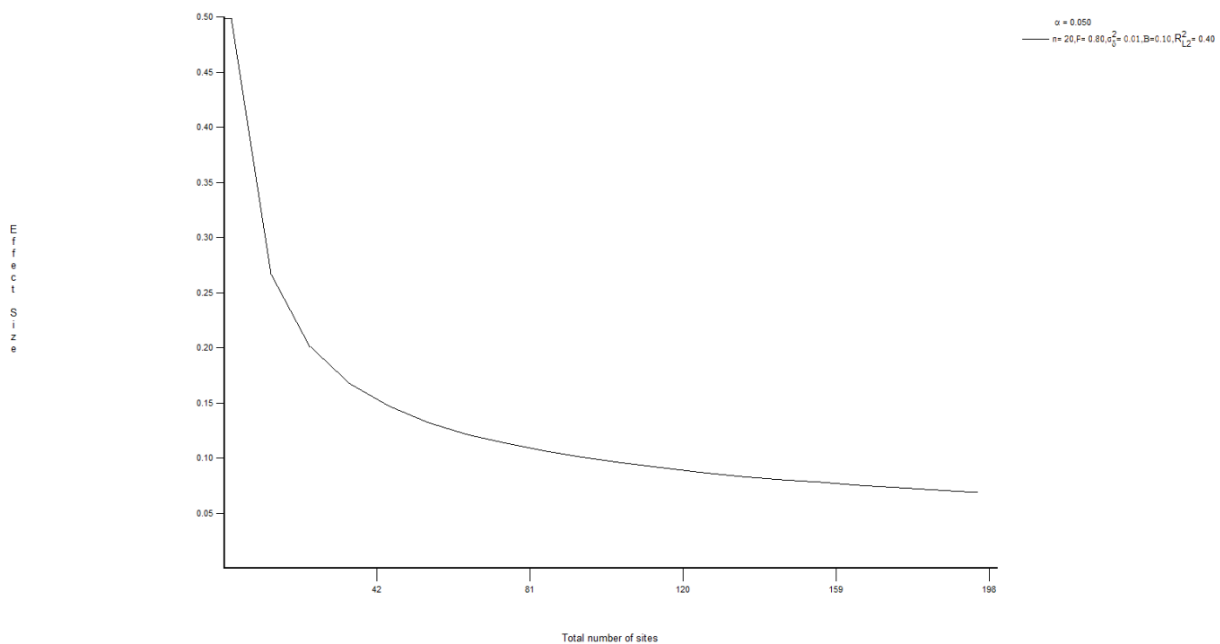


Figure 3 shows the full range of minimum detectable effect size by the number of schools surveyed. The number of clusters required to detect an effect size of 0.10SD and 0.08SD are 100 and 145 respectively. We expect the intervention to have a relatively small impact. The original study of York et al. (2019) found an effect of 0.11SD on a literacy test, while the study by Doss et al. (2022) found an effect size on a maths test of 0.16 among girls. An expected effect size of 0.10SD appears to be a reasonable assumption. However, we decided to further reduce the expected effect size to account for the potential contamination of the control group. The efficacy trial showed that at least 15% of parents in the treatment group shared the test messages with parents belonging to the control group. If the messages shared have the same impact of the messages received directly, we should expect an impact of the intervention of 0.015SD in the control group (for an estimated population impact of 0.10SD), and therefore an estimated impact of the intervention in the treatment group of only 0.085SD. Hence our power calculations assume a more conservative impact of 0.08SD to allow for a larger effect of contamination.

This set of assumptions delivers a required number of 145 schools for the study. The size is relatively large for an individual-level randomised study. However, the assumptions made above for the calculations are within the parameters usually found in the literature and we made provision for the risk of a small effect size, a high attrition rate, and a large volume of contamination of the control group.

The number of schools required to observe the same effect size on children beneficiary of Free School Meals (FSM) is considerably larger. Assuming the number of FSM children in the study were in the same proportion found in the average population (25.9%), implying an average cluster size of 5 FSM children per school, the MDES would be equal to 0.16.

At the time of writing this protocol the recruitment data suggest that 33% of children in our sample are FSM beneficiaries – a larger percentage than the one observed in the UK population (25.9%) because of the larger number of schools from disadvantaged areas

included in our study. This implies about 7 FSM children per class per school out of the 20 that we expect will be part of the trial after accounting for attrition. Keeping all other assumptions the same as in the power calculations for the overall sample, this would result in an MDES for this subgroup of 0.13 with 145 clusters (third column of Table 3).

This value is considerably larger than the expected value in the overall population (0.08). However, it should be observed that the same intervention was found to be much more effective among children performing poorly at baseline. For example, York et al. (2019) found that children under the median of the outcome distribution at the baseline experienced an impact of the intervention of about 0.30SD. Poor baseline performance is not the same as social disadvantage, but the two are correlated and it is reasonable to expect that the intervention will have a larger impact on disadvantaged children, in such a way that an expected effect size of 0.13SD is well within our expectations for this sub-group.

## Randomisation

We will use complete and blocked random assignment. Randomisation will take place at the parent level. The randomisation is complete in the sense that 50% of children will be assigned to the intervention and the remaining 50% will be assigned to a control group (which will continue with “business as usual” with no intervention). All children in the selected class of the recruited schools will be either assigned to the intervention or to the control group, and class size will not affect assignment. Randomisation is blocked in the sense that it will be conducted within strata (the “blocks”), which share common characteristics. The blocks in our study are the schools. Blocking increases precision and statistical power to the extent that the blocks explain the variability in the data. For example, in the power calculations we assumed that the schools explain 10% of the variance in test scores. This occurs because observations in the blocks share some characteristics (for example the socioeconomic environment or the quality of the teaching staff) that are correlated with the outcome.

Randomisation will be performed as follows:

- Child-level observations will be assigned a random number
- Observations will be sorted by block (the school) and then by the random number within the block
- The first observation in the sorted dataset will be allocated to the treatment
- Each subsequent observation will be assigned to have the opposite allocation of the previous observation

The randomisation will be conducted by the evaluation team, using Stata 17. The randomisation process will be recorded in the syntax and log files used to carry out the randomisation. The results of the randomisation will be shared with the delivery team so that the programme can be delivered to parents assigned to the intervention. Analysts will not be blind to trial arm allocation<sup>3</sup>.

---

<sup>3</sup> The trial is conducted on a sub-sample of all pupils in Receptions years: in multi-class schools one class is randomly selected to take part in the trial. The randomisation procedure related to the trial is conducted by the evaluator which will select half pupils within the

## Statistical analysis

### Primary analysis

The primary analysis will estimate the impact of the intervention on children's language skills as measured by the LanguageScreen test. The estimation will be the difference-in-difference between the children assigned to the intervention group and the children assigned to the control group over time. The estimation will be conducted regardless of compliance. That is we will estimate the "intention-to-treat" (ITT) effect, representing the average effect of the intervention on the population.

We will estimate the ITT effect using an ANCOVA (analysis of covariance) model, in which the post assessment score is regressed against the pre-intervention assessment score as in the equation below:

$$y_{ist} = \beta_1 y_{ist-1} + \beta_2 T_{is} + \alpha_s + \varepsilon_{is}$$

where  $y_{ist}$  is the test score of child  $i$  in school  $s$  at the endline,  $\beta_1$  is the correlation between the scores before and after the intervention,  $T_{is}$  is treatment status for child  $i$  in school  $s$ ,  $\alpha_s$  are school-level fixed effects obtained by including 144 dummy variables for each school in the dataset except one, and  $\varepsilon_{is}$  is a child level error term. The standard errors are clustered at the school level to account for homogeneity of observations within schools.

As a robustness check we will also: a) extend the model above to include other baseline covariates (representing characteristics of the child, the parents, and the school) to increase precision and remove any difference between the project and control observations that was not eliminated by randomisation, and b) run a multi-level regression model to account for the hierarchical structure of the data (children within schools within regions).

### Secondary analysis

The same ANCOVA model employed in the primary analysis will also be employed to assess the impact of the intervention on the secondary outcomes: scores of the revised Head-Toes-Knees-Shoulders test. In addition, the study will assess the impact of the intervention on intermediary outcomes to better understand the mechanism of impact. Online surveys will be conducted with all parents (participating and not participating in the intervention) before and after the intervention. The surveys will include questions on parents' attitudes and behaviours towards parenting and towards the school. They will be asked the frequency of home learning activities, such as reading stories to their child or singing rhymes, and attitudes and about their degree of interaction with the school and school staff. Changes in parents' behaviour in the project and control group will be compared and they will help explain the effectiveness of the intervention.

---

selected class. However, the intervention will be delivered to half pupils, regardless of how many classes the school have. The randomisation procedure related to pupils not included in the trial is conducted by the delivery team.

## Estimation of effect sizes

The estimated impact on pupils' test scores will be converted into a Hedges'  $g$  effect size. The impact estimates obtained using the ANCOVA regression model will be converted into standardised effect sizes in order to facilitate the comparability of effects across studies. The effect size will be standardised by the unconditional pooled standard deviation of the treatment and control groups:

$$S_{within} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

In the case of LanguageScreen we might be able to standardise the effect size by a population-level measure of the standard deviation rather than by the pooled unconditional calculated using the sample data, because the firm providing the test has collected assessment data on a large sample of children across the country.

The effect sizes estimated using multi-level modelling discussed in the robustness check will be standardised in a different way. We will use a standard deviation based on the total variance to account for the nested structure of the data and for potential differences between schools, so as to avoid an overestimation of the effect size (EEF, 2022).

For completeness, we will also report effect sizes as *Hedges g*, by adjusting the standardised effect size by the factor:

$$J = \frac{3}{4df - 1}$$

where  $df$  are the degrees of freedom used in the calculation of the unconditional pooled standard deviation above (Borenstein *et al.*, 2009). In addition, we will provide in the report all the parameters used in the calculations (sample sizes and variances) so that readers will be able to compute alternative definitions of effect sizes.

Ninety-five per cent confidence intervals will be reported around the resulting effect size, calculated by inputting the lower and upper confidence limits for the coefficient on the treatment variable from the regression model into the effect size formula.

## Subgroup analyses

We will conduct three subgroup analyses: by FSM eligibility status, by EAL status, and by baseline pupils' skill levels.

As discussed, behavioural interventions in education have been found to be more effective for children at a socioeconomic disadvantage. Given that our study sample includes children aged 4-5, we will consider as disadvantaged children those children that are either eligible for the Early Years Pupil Premium or in receipt of Free School Meals. Data on EYPP and FSM status will be obtained through linkage to the NPD data.

There is some evidence that children learning English as a second language (EAL) are at some risk of literacy difficulties (August and Shanahan, 2006; Kieffer, 2008). It is possible that the programme could have significant benefits for this group. However, the intervention is

effectively delivered through parents and if they are non-native English speakers they might not engage effectively with the text messages. The expected impact for this group of children is therefore uncertain and worth investigating.

Previous studies of the same intervention have found it to be more effective for children that scored poorly at the baseline (York, Loeb and Doss, 2019). The intervention is expected to be more effective for children with lower baseline skills. We will assess the impact of the intervention along the baseline outcome distribution using a method that will be described in the SAP, and that will consist of assessing the impact of the studies separately for children above and below the median test score at the baseline as in the original study by York et al. (2019), or another method such as for example quintile regression.

In line with EEF Statistical Analysis Guidance, two approaches will be used and reported for the subgroup analysis. Firstly, a subgroup analysis will be conducted through estimating a separate model for the selected subgroup. Secondly, we will repeat the primary analysis specification but additionally include an indicator for EYPP eligibility and an interaction term for EYPP eligibility and treatment status.

### **Analysis in the presence of non-compliance**

The intervention consists of sending tips by text messages to parents. In order to be successful, the messages have to be delivered, they must be read by parents, and they must be acted upon. Our definition of compliance will be based in the first instance on message delivery.

The programme consists of a package of about 120 messages. In the efficacy trial, all messages were delivered to 73% of parents, and only 6% of parents received fewer than 50 messages (Stokes *et al.*, 2022). This information was obtained from delivery records made available by the delivery team and included: the number of messages sent, those successfully delivered, and those failed to deliver. The same information on the delivery of text messages will be made available for the current study. Failure to deliver messages can be the result of either withdrawal from the study or inability to reach a specific phone number because of changes in phone numbers. In the analysis plan we will develop a measure of compliance such as full compliance (receiving all messages) or partial compliance (receiving less than 50% of messages) and we will describe the Complier Average Causal Effect (CACE) analysis of the impact of the intervention on the “treated” pupils.

An endline survey of all parents will collect data on the use of the messages as reported by parents. Parents will be asked whether they read and understood the messages received and which activities they undertook with their children after reading the messages. This information would in principle allow to build a measure of compliance that reflects more closely the implementation of the intervention. However, the survey will be sent to all parents to be filled online and not all parents will be able to answer all the questions. In the efficacy trial less than 50% of parents provided responses to the online survey (Stokes *et al.*, 2022). If the parents’ survey delivers a sufficiently large number of responses, the information will be used to build a more accurate compliance measure. The efficacy trial obtained a response rate of 65% for the parents’ survey and we are aiming at a target of at least 70% for the current study. Lacking any information on whether parents understand the text messages correctly and on the intensity with which parents change the home learning environment after receiving the

messages, we tentatively define “compliers” as those parents that in the parents’ survey report changing their behaviour after reading the text messages in at least 50% of cases considered by the survey. If few parents respond to the survey, then the information will be used to understand the fidelity of the intervention and how successfully the intervention was implemented as described below in the section on the implementation and process evaluation.

### **Additional analyses and robustness checks**

As mentioned in the primary analysis section we will conduct the following robustness analyses: a) we will extend the ANCOVA model to include school-level and child-level baseline covariates to increase precision and to remove any difference between the project and control observations that was not eliminated by randomisation, and b) we will run a multi-level regression model to account for the hierarchical structure of the data (children within schools within regions).

We will also conduct an analysis accounting for contamination of the control group. The evaluation of the efficacy trial found some spill-over effects in the parents’ survey (Stokes *et al.*, 2022). Parents receiving the intervention reported sharing text messages with other parents in the following proportions: 22% (150 parents) sharing with parents of children in other schools; 11% (72 parents) sharing with parents of children in same school but different year; 21% (141 parents) sharing with parents of children in same school and year; and 47% (318 parents) not sharing. Worryingly, up to a fifth of parents shared text messages with parents of children attending the same school and year, and therefore potentially with parents of children included in the control group. If parents in the control group are receiving text messages, the impact of the intervention can be underestimated.

During their interviews with the evaluation team some parents suggested that COVID-19 may have led them to increase message sharing, whereas other parents suggested that COVID-19 may have led them to decrease sharing, so that the overall impact of COVID-19 on message sharing is unclear. Some parents suggested that the challenging learning environment during COVID-19 led them to share any help that became available including text messages, while others suggested that they might have had more opportunities for sharing if they had been able to meet in person. As for the frequency and consistency of sharing, some parents reported sharing all messages, while others reported sharing only some of them.

In the surveys, all parents in the treatment group will be asked whether they shared the messages received, with whom and how often. Similarly, parents in the control group will be asked whether they received the Tips by Text messages from other parents and how often. This information will be used to build a continuous measure of treatment intensity in the control group, which will be used to characterise subjects in the control group as “contaminated” or “treated”.<sup>4</sup> For example, parents in the control group who received one or two messages and reported not using them, will not be considered as being “contaminated” or “treated”, while parents in the control group receiving and using the messages will be classified as “contaminated” and “treated”. Provided the information is available from a sufficiently large

---

<sup>4</sup> See Sussman *et al.* (2010) for a discussion of contamination framed as a compliance issue and Imbens and Rubin (2015) for a discussion of the method of 2-sided compliance discussed above to address contamination.

number of parents, it will be used in the analysis in two ways that we will be described in more detail in the analysis plan:

- To model bounds of effects around the estimated project effect. For example, if the project impact is 0.10 of a standard deviation and we document consistent sharing of messages with 10% of parents in the control group, we can infer that the intervention effect is underestimated at most by 0.01 sd. The main result reported by the study will be the observed effect ignoring contamination. The bound above the observed effect will represent the range of effects under different hypotheses on the extent of contamination. The observed effect should be interpreted as being the effect “at least”, while the upper bound accounting for the estimated contamination can be interpreted as the effect “at most”.
- To estimate the average effect of the intervention on the treated children (ATT) whether reached directly or indirectly via sharing. This can be done with the estimation of a 2-sided complier average causal effect model (CACE).

An obvious way to prevent contamination is asking parents not to share text messages. We initially considered the opportunity of sending additional text messages to parents during the intervention or at the beginning of the intervention with the purpose of preventing parents from sharing tips with other parents. However, it was observed that such messages could lead parents to behaviours that would not replicate under non-experimental conditions. It is impossible to predict the direction and the size of the behavioural impact of further messaging since messages could either motivate or demotivate parents from adhering to the intervention. In addition, multiple messages sent by the delivery team and by the evaluation team may confuse parents and lead to withdrawals. This was also confirmed by an external reviewer explicitly consulted on this matter. However, some amendments were made to recruitment documents including statements in the Parent information sheet to not share messages. In the MOU and in the School Information sheet, clarity and roles were added around not sharing texts.

### **Missing data analysis**

Missingness is most likely to arise through non-completion of assessments at both the baseline and endline. We will document both the extent and pattern of missingness.

First, we will calculate and report the number of complete observations, and we will illustrate how the observations were lost using an illustrative flow diagram. Second, we will compare missing observations between the treatment and the control group to detect the presence of differences in the frequency of missing values between the intervention and the treatment group that could bias the analysis. Third, we will explore the determinants of missing values when the proportion of missing observations in the data is sufficiently large.

More specifically, we will analyse the determinants of missing observations when these account for more than 5% of the data in the sample. If there is less than 5% missing data, then it is unlikely that missing data will result in bias in the impact estimates (Schafer, 1999). We will investigate the determinants of missing observations, such as pre-intervention characteristics of the child and of the school. To identify the determinants of missing observation we will use a logistic regression model, using any missing observation as the dependent variable and baseline characteristics as explanatory variables. The model will include baseline values of child and parents' characteristics (such as gender and other

demographics), and school characteristics (such as location, area-level social disadvantage, and practitioners' characteristics). The model will be separately run in the intervention and in the control group to verify if the determinants of missingness differ in the two groups.

The approach to missing data analysis will be specified in more detail in the SAP. We anticipate that any such analysis will be focused on the primary outcome, in line with EEF analysis guidance.

### **Statistical analysis**

Quantitative analysis of the parent survey at both baseline and endline will include basic statistical tests to identify changes in practice or behaviour, using descriptive frequencies and cross-tabs, as well as inferential statistics. Data will be analysed to compare intervention and control groups and controlling for change over time (if we have sufficient repeated measures data). This would be achieved using an ANCOVA model or equivalent non-parametric design. Further details will be provided in the SAP.

These findings would then be triangulated with findings from qualitative and monitoring data to explore what factors may contribute to the findings or highlight areas for closer consideration where findings from different sources presenting conflicting pictures.

### **Delivery information**

Programme information provided by the delivery team would be used to provide broader insights into delivery that can be used to enrich other analyses. The delivery team will be asked to provide information about the current price of the intervention for schools/individuals outside of the trial, whether they have made any changes to the intervention during delivery that would mean that the price would be different for the intervention as it has been delivered during the trial, whether they anticipate any change to the price for future delivery, and whether they anticipate any change in the price of the intervention if it were to be delivered on a larger scale. Further detail is provided in the cost evaluation section of the protocol.

## **Implementation and process evaluation (IPE) design**

### **Research questions**

The aim of the IPE is to establish fidelity in implementation, identify the factors influencing impact and contextualise the quantitative findings. We would also explore programme effectiveness, perceived impact of different elements of the intervention and considerations for a wider roll-out in line with [EEF guidance](#) on IPEs for effectiveness trials. It will also be important to build on the existing learning gained from the efficacy trial.

The main findings from the IPE in the efficacy trial were:

1. **The intervention was implemented well by parents, and they felt it was acceptable and useful.** It is important to explore how well the programme was implemented as it can explain the findings from the impact strand of the trial. We plan to explore this again here because the efficacy trial was conducted during the global pandemic when parents had more time at home with their children. We do not know for sure what impact the school closures had on the implementation of the programme.

The effectiveness trial will mean testing can take place in a context which is not affected by a global health crisis and enable us to explore whether the intervention remains acceptable and useful to parents, and that they still implement it as intended. This relates to the dimensions of dosage, reach, responsiveness and, perceived impact.

- 2. There were signs of potential contamination between the intervention and control group parents.**, Parents receiving the intervention reported sharing text messages with other parents in the following proportions: 22% (150 parents) sharing with parents of children in other schools; 11% (72 parents) sharing with parents of children in same school but different year; 21% (141 parents) sharing with parents of children in same school and year; and 47% (318 parents) not sharing. So, around a fifth of parents reported sharing text messages with parents of children attending the same school and year, and therefore potentially with parents of children included in the control group. In-depth qualitative interviews with 25 parents suggested that the impact of COVID-19 on sharing was ambiguous. Some parents suggested that the challenging learning environment during COVID led them to share any help that became available including text messages, while others suggested that they might have had more opportunities for sharing if they had been able to meet in person. Due to the individual randomisation of the trial participants and the possibility of intervention and control parents knowing one another, it will be extremely important for us to measure this to understand any spillover effects for the control group. The IPE has been designed to enable us to assess the occurrence of text sharing in-depth, by asking parents to provide additional details such as how often this occurred, the reasons for it and what happened as a result of sharing. Included in the methods are surveys with parents in the control and intervention groups at baseline and endline. The IPE also includes in depth-interviews with two sets of parents, one longitudinal group and another just at the endline. The longitudinal group will be interviewed at two time-points, the first will at the mid-point of the intervention (in Spring 2025), when parents should have a good recollection of their behaviour at that time and, in a follow-up interview, towards the end of the intervention. The efficacy trial only included interviews with parents at the end of the intervention, which is not as reliable as it relies heavily on parental recall, which does not always yield accurate responses months after an event. Designing the IPE to include a larger number of parents at multiple time-points should allow a more thorough analysis of changes to parents' usual practice than was possible in the previous trial. This relates to the dimensions of dosage, reach, monitoring of control group, and perceived impact.

It is important to note that one of the aims of effectiveness trial IPEs, as set out in EEF guidance, is to capture information about any changes made to the programme following an efficacy trial. However, in this case, only minor changes have been made to the programme itself. We understand these to be a change in the provider sending the texts; these will now be sent by FireText. There are also some minor changes to some of the texts. Further, to allow adequate time for the assessments, there will also be a minor change in dosage as the delivery window will be 1-2 weeks shorter than the original trial.

Using a mixed-method design, the IPE will explore the following dimensions specified in the EEF IPE guidance (2022): fidelity, adaptation, dosage, quality, reach, responsiveness,

perceived impact, context/moderators, programme differentiation, monitoring of control group, cost and any possible mediators. We present the IPE research questions related to these dimensions under the broad headings of fidelity, implementation and outcomes. A full mapping of these is shown in Table 4, however in the following section we briefly mention the measures we will use to address each Research Question. The subsequent section on Research methods also includes more detail of these measures.

**Fidelity.** To establish fidelity, defined as the extent to which the implementers adhere to the treatment model, we will look particularly at whether the programme is delivered as intended, with compliance to the planned approach by the delivery team, and how well this has been achieved. This will be explored using the following research questions:

- RQ 1: To what extent was the intervention implemented as intended? Did the delivery team send 3 texts per week? (Dimensions: fidelity). Measures include – data from BIT on the delivery of text messages. In the previous trial, delivery records were available for all parents (and thus pupils) in the treatment group. These provided information on the total number of messages sent, the number successfully delivered, and the number which failed. Information was also provided on the number of delivered messages that had been re-sent. As stated above, minor changes have been made to the programme for this trial We plan to access this data in the same way again to explore the dimension of fidelity.
- RQ 2: Does the intervention remain acceptable to parents, and do they find it useful (in terms of the short-term outcomes specified in the Theory of Change), for example, do they feel it adds value to the home learning environment, promotes or improves parent-child interactions in the home or improves home-school relationships. (Dimensions, responsiveness, perceived impact). Measures include parent surveys and parent interviews. The efficacy trial established that the intervention was acceptable, and parents found the texts useful. However, this was run in the context of a global pandemic with schools locked down, so we will explore it again in the current context, where children are attending schools as usual and there is far less homeschooling happening.

**Implementation.** We will consider the wider range of issues which affect implementation, including the necessary conditions for success and barriers to successful implementation, using the following questions:

- RQ 3: What are parents doing after receiving the texts and how different is this to their usual practice (what are the signs of the intermediate outcomes being achieved and thereby behaviour change)? This includes exploring the proportion of parents acting on the texts, activities that happened at home after receiving the texts, and how these activities were carried out, and whether engagement changed over time? (Dimensions: dosage, quality, reach, responsiveness, programme differentiation). Measures include parent survey and parent interviews. The efficacy trial endline survey found that the majority of parents put some (60%) or all/most (38%) of the texts into use after receiving them and this was backed up in the interviews with parents. However, we only had endline data in the efficacy trial, so this trial will enable an increased understanding of differences made to parent behaviour in the home-learning

environment (HLE). This will also address the quality dimension as the measures will enable a review of the content of texts and identification of any differences between the three different types of texts and what parents do with them.

- RQ 4: What happened in the control group? Specifically, was there any contamination of the control group by treatment group parents sharing text messages. What was Business as Usual (BaU), including in terms of the home learning environment, parental engagement and existing texting interventions and platforms? (Dimensions: monitoring of control group, programme differentiation).  
Measures include parent survey and parent interviews. As described above, the efficacy trial found that parents in the treatment group had shared text messages with other parents. The IPE will be able to obtain further information about the occurrence and impact of text sharing this time, as parents in the control group will be surveyed at baseline and endline.
- RQ 5: What are the facilitators/barriers to the programme? Do any facilitators or barriers apply specifically to parents/pupils from disadvantaged backgrounds? (Dimensions: context/moderators, mediators contained in the intervention theory of change). Measures include parent survey and parent interviews. The IPE in the efficacy trial was not designed to analyse barriers faced specifically to parents or pupils from disadvantaged backgrounds. This IPE will allow analysis by free-school meal status as schools will be asked to provide this information. In addition, the parental surveys will ask demographic details which can be used in the analysis. The depth-interviews will also include detailed questions on barriers faced by parents including probes on language, technology and timing needs, which will enable a more nuanced analysis.
- RQ 6: Was the intervention an acceptable investment for schools from the perspective of management (considering any staff time input required, for instance to update phone numbers)? (Dimension: cost). Measures include a short data collection tool sent as a survey via Snap Surveys, as it was found in the efficacy trial that involvement in the programme from a school's perspective is minimal and generally only involves staff time required to update parent phone numbers. The efficacy trial involved interviews with a small number of schools' senior leadership teams in case study schools and covered a larger range of issues. This IPE has been designed to capture data from a larger number of schools and purely on the cost implications (actual and potential) of their involvement in the programme. A short research tool will be designed, to be used at endline only, to capture data on any costs incurred by schools and views on what costs would be acceptable to them in a real-life future scenario (either if they ran it themselves or it was run by an external organisation).
- RQ 7: What do the findings suggest for how the intervention could be rolled out at a larger scale? Would anything need to be adapted for the intervention to be deliverable at a larger scale? Could the intervention be scaled and become more targeted at disadvantaged children? (Dimensions: fidelity, adaptation). Measures include analysis of findings on acceptability through contacting schools. Also link to cost findings from cost analysis (carried out by the Impact Evaluation team) of schools and/or securing another provider. Also include any information gathered by the delivery team.

**Outcomes.** It is important to explore parents' perceptions of the programme at two time points: pre-programme (covering anticipated effects) and post-programme (covering effects based on experience). Their pre-programme perceptions will in likelihood affect their commitment, and therefore programme impact and effectiveness. Here we will explore the following overarching research question:

- RQ 8 How did parents perceive and experience the effects of the programme, exploring pre-programme perceptions of anticipated effects and post-programme perceptions of actual effect? (Dimension: perceived impact)
  - a) Did parents perceive that the intervention led to higher language and numeracy child outcomes?
  - b) Did parents perceive that the intervention led to improved self-regulation of their child?
  - c) Did parents receiving the intervention report that they engage more with the school than they did pre-intervention?
  - d) Did parents receiving the intervention report that it improved parent–child relationships?
  - e) Is the intervention perceived by parents to improve the home learning environment?
  - f) Did parents prefer one area of support (for example, language) over another (for example, social development or numeracy)?
  - g) Were there any unintended or negative effects of the intervention? (for example, parental views about the texts leading to disengagement, stress and pressure from expectations and comparisons) Does the intervention displace/constrain other activities which might occur in the HLE? (e.g. replacing other behaviours with the new ones).
  - h) To what extent does the intervention make a difference to disadvantaged parents in contrast to those who are more advantaged in terms of responsiveness, perceived impact, and change from existing practices? (Dimensions: context/moderators, mediators, reach).

Measures include surveys and interviews with parents. These will be included as measures in the survey of parents and questions in the interviews. This trial will enable us to ask these questions in a non-pandemic context, when children are back to 'normal' activities related to school, family/friends and wider extra-curricular activities such as home-learning, after-school clubs, sports activities etc which were not possible during lockdowns. On the final outcome regarding disadvantage, depending on survey response rates, we may have potential for some subgroup analysis. The survey could be descriptively assessed for any differences between advantaged and disadvantaged (based on the FSM variable). These issues will also be explored in the interviews and analyses done (on FSM). However, given known disparities in smartphone ownership and contract types, it may be more difficult to engage more disadvantaged parents, which may affect our ability to lead robust analysis on this theme. The

counter-balance comes around a cheap, scale-able universal intervention that might benefit more disadvantaged parents who take part. More targeted (but costlier) interventions might lift more load but would be less scalable.

## Research methods

Drawing on the EEF IPE guidance, (Humphrey et al., 2016, EEF, 2022) as well as the learnings from the efficacy trial, we propose a pre/post, multi-component design, based around a triangulation of mixed methods, to examine the research questions above. To do this we propose:

- Survey of parents (baseline and endline).
- In-depth interviews with parents, including longitudinal interviews with a cohort of parents (halfway through and at the end of intervention).
- School endline data collection.
- Review of texts sent.
- Analysis of any data collected by delivery team such as responses to texts or feedback surveys.

### **Survey of parents – baseline**

In line with the IPE guidance from EEF 2022, which states that all trials will capture usual practice at baseline in all trial arms, we have included a baseline survey for all parents. The baseline survey of all parents in the trial will take place prior to randomisation, in late September and October 2024. The baseline survey will focus on determining usual practice in the home learning environment (HLE) and parental engagement with their child's school. The survey will include questions on parental attitudes and practices around home learning, their confidence in supporting their child's learning related to the school curriculum in the HLE and engagement with the school.

Questions will be designed so that they can be repeated in the endpoint survey. This will enable us to measure change over time as well as change between groups, including quantitative measures of certain activities they do with their child and the frequency and duration of such activities. This is in line with the IPE guidance which states that, where possible, evaluators will include a quantitative measure of usual practice (e.g. times per week, or number of hours doing an activity) so that this can be analysed to assess any changes to practice that have occurred over time in the control versus treatment groups.

The baseline survey will also explore other usual practices by asking parents about any text-based interventions they've been involved in/are currently in at the time of the survey.

### **Survey of parents – endline**

The endline survey of parents in the treatment group will take place towards the end of the summer term (June and July 2025 - further details below). To enable comparisons across the two time points, questions around parental experiences of learning at home, the range, frequency and duration of activities done in the HLE and engagement with their child's school will be repeated.

The endline survey will also gather parental experiences and views of the text messages, including how often they received the texts, how interesting and useful parents found the texts

and, what they specifically did after receiving the three different types of messages (Fact, Tips and Growth). Parents will also be asked questions around the appropriateness of the type, frequency, duration, and timing of the messages, and whether they would like any changes made to the intervention. To measure the extent of spillover, due to parents sharing the texts with other parents in their networks, parents will be asked if they have shared the messages, and if so, how often, to whom (i.e. parents in their child’s class, other classes in the school, other schools, family etc), and their motivations for doing so. It will also include questions around which activities parents engaged in at home with their child after receiving the texts, and which outcomes and behaviours changed in them **and** their child such as engagement, enjoyment and confidence in carrying out activities in the HLE and perceived changes in language and self-regulation. The survey will also ask about any barriers to implementation, and reasons for non-compliance. We will also explore any unintended consequences of the intervention, such as other behaviours being displaced or negative experiences.

The endline survey of parents in the control group will take place at the same time as the treatment group (June and July 2025) and will include the same questions on parental attitudes and practices around home learning and engagement with the school. As with the treatment group, it will repeat the same questions around frequency of home-based activities as the baseline survey, so that we can make comparisons across time as well as across groups. The endline survey of parents in the control group will also explore the extent of contamination, if parents in the control group received texts forwarded on by or copied from the treatment group (or otherwise). If this is found to be the case, they will be asked for more detail, including the frequency this occurred, the types of messages they received (i.e. Fact, Tip or Growth) what they did after receiving the messages and about any perceived impacts.

Below is a table based on a template from the IPE guidance which sets out the main broad outcomes we wish to capture with the surveys.

Table 4: Broad outcomes of parent surveys (treatment and control)

	Baseline survey	Post intervention survey
Control group	Programme differentiation –Capturing baseline frequency and duration of home learning activities. Usual practice – are any other text based or tips-based interventions currently running	Capturing endpoint frequency and duration of home learning activities  Contamination – Were parents forwarded the texts, if so, how often, perceived impact of receiving, and did behaviour change.

Treatment group	Programme differentiation –Capturing baseline frequency and duration of home learning activities. Usual practice – are any other text based or tips-based interventions currently running	Capturing endpoint frequency and duration of home learning activities  Experience and views on receiving the texts  Unintended consequences – were there any negative consequences of receiving the texts.  Contamination – did the treatment group forward the texts to the control group, and if so, how often/to whom/reasons.
-----------------	---	---

The design of the surveys will be led by the IPE team at IES, and will draw on the survey in the efficacy trial, including learnings from the previous process and covering the key issues identified for the effectiveness trial. The design of the surveys will be closely aligned with the theory of change, ensuring there are adequate questions around contamination and around parental and child behaviour change.

The surveys will involve a limited number of short, well-structured questions with clear rating scales, which will reduce parental burden and help to maximise response. For context, the endline survey in the previous trial was 18 questions long and took 10 minutes to complete. However, this did contain some questions relating to the COVID-19 pandemic which can now be removed. We envisage the surveys to be of a similar length to previous as this achieved a good response rate however, this will be finalised when the design is complete.

In terms of design, we will ensure key aspects of the protocol are agreed before designing specifics of the surveys so that time is spent economically. Design will begin in June 2024, with a review phase taking place in July 2024. The final design will be ready for programming into the Snap Survey software early August 2024. The survey will be subject to rigorous testing by the evaluation team from mid-August 2024. Once any changes have been made following testing, the survey will be ready to go live in late September 2024.

The IPE team at IES will administer the survey, sending invites to parents and hosting on the survey platform. Survey links will be individualised so that reminders can be sent to parents who have not yet responded. Reminders will provide behavioural incentives to encourage survey completion – such as the reasons why it is important and how this evidence could make a difference. The proposed plan is to send one invite text and up to two reminder texts at each survey wave. To ensure data quality and reliability, a response rate of between 50% and 60% would be desirable. Data will be collected on the survey platform and analysed by IES. See below for further detail, discussion, and debate of this proposed process.

### *Delivery of surveys to parents*

The surveys of parents will again be developed with online platform Snap Surveys, which allows completion on mobile devices via hyperlink or QR code and are smartphone/smart device compatible. At the start of the 2024/2025 academic year, before randomisation, parents' mobile phone numbers and email addresses will be collected as part of the wider baseline data collection exercise. NIESR collect these data and share them with IES for the purpose of contacting parents for surveys and interviews, in order to deliver the IPE side of the evaluation. As discussed above, the endline surveys will take place later in the summer term of 2025, after the majority of the intervention has been completed, enabling parents to reflect on the process. Given that the intervention runs until the very end of the school year, there is a balance between placing this as near to the end of the intervention as possible and giving parents time within the school term to respond. Where parents have dropped out, it will not be possible to track change over time, but it will still be possible to make comparisons at single time points between practice amongst treatment and control parents.

If there are 145 schools in the trial with an average of 26 pupils per school taking part then this means we have a pool of around 3,770 parents to survey. All of the parents who have not withdrawn from the study and have provided contact details will be sent the survey to complete. In the previous efficacy trial, we had planned to send out the survey invites by email but found that many parents' emails were missing from the school data templates. Additionally, only treatment group parents were surveyed in the efficacy trial, and only at endline. Since the delivery team already had phone numbers for the treatment group parents, they agreed to send the surveys and associated reminders to parents via SMS.

In this trial, in line with the updated 2022 EEF IPE guidance, we are proposing a baseline and endline survey of both control and treatment groups. This is a difference and addition from the previous trial. The baseline survey will occur before randomisation and before any phone numbers of the treatment group parents are shared with the delivery team. In addition, we are also surveying the control group and currently there is no plan for the delivery team to have control parents' phone numbers at any point during the trial. Thus, for the control groups' baseline survey and endline survey, the delivery team are not in a position to send out the invites under the current data sharing plan and agreement. Therefore, personalised survey links will be created for each parent who has not withdrawn from the evaluation. We are happy for the delivery team to co-develop the wording of the invite sent with the survey. Having the evaluation team (IES) sending the invites out will also preserve the independence of the evaluation. The survey will be open for approximately five weeks, with invites sent out in late September 2024 (dependent on timely receipt of parent contact details from schools) and the survey left open until the day before the texts start being sent to parents (by the delivery team) currently November 11<sup>th</sup> 2024. The survey will be programmed to include two reminders; the first will be two weeks after the survey has gone out and the second will be one week before the survey is closed. As survey links will be individualised, reminders can also be individualised, to ensure that only parents who have not yet responded are contacted. This could not be achieved if the delivery team launches the survey, as it would be unethical for them to receive information on the parents who have responded to the evaluation request.

The delivery team have raised concerns about this process, and we address these concerns below. The delivery team have a reservation about evaluators contacting research

participants. They have a worry that a text message invite to take part in a survey from the evaluators will lead to parents wishing to drop out of the trial, and that a text message from a different organisation other than the delivery team could be confusing. Given that contacting parents is necessary for the evaluation activities to occur, we have to carefully consider how best to achieve this without causing harm to the trial or intervention.

The delivery team have offered to contact research participants on behalf of the evaluation as they feel that this would be less confusing for participants to receive texts from one organisation rather than two. However, currently there is no plan for the delivery team to hold parents' phone numbers at the baseline survey point, prior to randomisation, and there is also no plan for them to hold the control group phone numbers. This was to preserve data protection, and to only share what was necessary.

The delivery team also queried how IES would deal with a withdrawal of a parent if it came to them via text message. If this were to happen, IES would immediately inform both BIT and NIESR and the parent would be marked as withdrawn and not included further in the trial. As a research organisation, fully compliant with GDPR duties, and we are experienced and used to running research in schools, both with parents and young children, and we often handle the consent and withdrawal phase. We operate on an informed and ongoing consent basis, where participants are free to withdraw from research at any point without suffering any negative consequences. Participants who withdraw from the trial are respected and not contacted further.

Finally, there was a concern about how the evaluators would deal with a safeguarding issue if it came to them directly. It would depend on the nature and specifics of the report but if we were alerted to a risk of harm, we would support an individual to identify support within their own networks. If a child were involved, we would contact their teacher and find out who to disclose to, to ensure they got support. The evaluation team can create a protocol for the fieldwork team, to ensure everyone is aware of the process. It is very rare for evaluation teams to need to make a disclosure, but if a vulnerable adult/parent were a risk to themselves or others (including their children), we can involve emergency services. We would also keep evaluation partners and the delivery team up to date with any such report or issue.

The proposal for us as evaluators to send invites and reminders at appropriate times to parents is based on usual practice of evaluation teams carrying out the activities as independent evaluators. Taking all of the above points into consideration, it is proposed that the evaluation team will send out surveys to parents, and the delivery team will input into the wording of invites, and the evaluation and delivery teams will alert each other of any withdrawals.

#### *Parent survey response rates*

The previous trial achieved a reasonably high response rate in the endline survey of parents (N=867, 65%)<sup>5</sup>. For the baseline survey with parents in this trial, we would anticipate receiving around the same response rate, especially given the evaluation team will be contacting all the parents who have not withdrawn and may be motivated to engage in the intervention and

---

<sup>5</sup> At randomisation there were 1,325 pupils, so this signifies an estimated response rate of 65%.

associated evaluation activities. If funding is confirmed, emulating practice in the efficacy trial, a small prize draw will be offered to incentivise parent engagement. At the endline survey, we anticipate a slightly lower response rate (around 50% approximately) partly due to intervention and evaluation fatigue, plus a small number of parents may not understand why we are carrying out an endline survey. Again, if funding is confirmed, we will offer a small prize draw as an incentive. At both time points, we will use personalised reminders to increase the response rate. As stated above the approach will be an initial invite and up to two personalised reminders.

For both parent surveys we anticipate a good response rate across the four regions – North-West, South East, Yorks and Humber, and London.

### **In-depth interviews with parents**

In-depth interviews with parents in the treatment group will explore their perspectives and experiences of the intervention. The interview script will be designed by IES, with input from NIESR, and the interviews will be carried out by a joint qualitative team with members from each organisation who will work together on all aspects of the interviews. These interviews will be designed to explore similar issues as the surveys (and aligned to the Research Questions), but will provide richer qualitative data on experiences, views, and perceived impacts, as well as enabling us to explore unidentified themes. Parents will also be asked about the time it took them to participate in the intervention and in any associated activities in the HLE. The interview will also include questions on whether any costs were incurred and whether they thought these were worthwhile, for their child's learning or for the HLE. The efficacy trial showed these interviews were an excellent way of establishing rapport and capturing parents' honest views.

We propose a total of 45 interviews, with 35 parents. Firstly, we suggest conducting interviews with the same number of parents as the efficacy trial and at the same time point (25 interviews in June/July 2025) as this enabled us to understand and compare the experiences of different parents and pupils. In addition, we are also planning to conduct longitudinal interviews with an additional 10 parents, involving a mid-point interview (in February/March 2025) and a follow-up interview with the same parent towards the end of the programme (June/July 2025). These longitudinal in-depth interviews (10 parents taking part in two interviews so 20 interviews in total) will allow us to capture data while the intervention is ongoing, including any issues with fidelity, implementation, and spillover. Although this may seem a relatively high number of participants for qualitative research, this number will enable some analysis across regions (and possibly between parents from different demographic groups for example Free School Meal status (if these data are available at the time of analysis, see the section below on sampling). The addition of a longitudinal cohort will also enable us to understand their experiences over time.

All interviews will explore perceptions and impact of the different types of messages (i.e., Fact, Tips and Growth), using parents' own terminology to describe the texts, and providing examples of texts as prompts to ensure that participants are able to recall the specific elements of the intervention. Where possible, the interviews will be conducted online (to allow screen-sharing to use examples of texts as prompts). We will also give parents the opportunity to do the interviews by phone, to maximise participation and the diversity of the sample (parents will be sent examples of texts before-hand to discuss in the interview). The interviews will last 30-

45 minutes and will be organised at a time convenient for the parents. All parents will be paid £40 per interview as an incentive to participate and as compensation for their time, paid as a shopping voucher. Previously, we used £20 incentives in the efficacy trial, but found a higher incentive would benefit our sampling. The proposed figure of £40 is fairly typical for 45-minute interviews with members of the public.

The sample of parents for in-depth interviews will be drawn randomly from the information supplied by schools. We will split the sample according to the region they are in (the survey will include a question on region with a drop-down, parents will also be asked to write in their school's name). Within each region, we will randomly select parents to invite for interview. This will be done in proportion to the overall regional split i.e. there are more schools in the North-West and less in Yorkshire and the Humber so it would be appropriate to sample the parents according to the proportion of schools. These measures will ensure that we obtain as varied a sample as possible, including for parents/pupils from disadvantaged backgrounds (based on FSM status). Parents who have withdrawn from the evaluation process will be excluded from the sample. The baseline survey will also include an opportunity for them to opt-out of the interview.

### **School endline data collection survey**

The endline data collection with schools will be designed to gather information on any costs that were incurred during the intervention to the school. In the previous trial, these questions were only asked in the interviews with members of the Senior Leadership team. However, as there are no staff interviews in this trial, these data will be collected at endline only, in the form of a survey sent to school leads. They will be asked to complete this themselves or pass it onto the relevant staff member i.e. finance/school bursar to complete. We will also collect school demographics as part of this process. Please see the section below on costs for more details.

**A review of programme texts** will be helpful in exploring in more detail the content and number of messages that have been sent to parents and when, and whether this is associated with parental engagement. This will be an important part of monitoring delivery for dosage. During the efficacy trial, this review was limited as the platform could indicate whether texts had been delivered, but beyond delivery the platform cannot tell if a text has been read or not. If it is possible to collect additional data on this, this would be included in the review. Otherwise, the review will be limited to number of messages sent and delivered. As it stands this data can speak to dosage – how much of the intervention has been delivered, but not to reach or responsiveness – the rate and scope of participation or participant engagement. As discussed above in the Research Questions section, the parental survey and interviews will be the best source of information on these dimensions as it will ask for details on receiving the receipt of texts and what they did following delivery.

Finally, we would welcome **sharing of any relevant data collected by the delivery team** during the project such as any responses to texts, requests for extra support or feedback from families or school staff if collected. These would be analysed in light of the purpose of its collection, and data would be triangulated with the main evaluation findings. The delivery team would also inform the evaluation team if any schools or parents decide to withdraw from the programme, in which case they would be contacted by the evaluation team and asked about

their reasons for doing so. We would also collect any cost data from the delivery team to calculate costs in line with EEF cost evaluation guidance.

## Analysis

Following EEF guidance, the IPE data will be analysed in isolation of the impact evaluation data. Qualitative and quantitative data from the interviews, surveys and delivery information will be integrated in the final IPE analysis to create a rich picture of delivery, participant experiences and school views to investigate mechanisms of change set out in the Theory of Change model and inform impact analysis. The different research methods will capture different perspectives and different levels of detail.

Survey and delivery data will provide quantitative insight on overall trends and patterns across participating parents and schools, whereas qualitative interview data will provide in-depth information regarding a limited number of schools/individuals that can offer insights into what factors may contribute to the broader trends observed. The theory of change, alongside the analytic framework shown in Table 5, will guide analysis and synthesis for each IPE dimension. Information from different sources will be compared to identify where findings are supported across multiple sources, and where differing findings indicate that closer analysis is needed to understand minority perspectives/experiences, differences in perspective/experience by stakeholder, or differences between reported observed/measured experience. These findings and data will be synthesised to create robust and nuanced conclusions that test mechanisms of change set out in the Theory of Change model and inform impact analysis.

It should be noted that, in terms of the surveys, anything over 500 responses will enable subgroup analysis. However, the parent surveys will only be able to collect data on perceived impact from a parental perspective.

### **Qualitative data analysis**

Interviews will be digitally recorded with the agreement of participants and transcribed verbatim. Data will be analysed using a 'framework' approach, from an analysis of interview transcripts and other materials collected by evaluation and delivery teams. Framework is an excel-based qualitative analysis tool that ensures that the analytical process and interpretations from it are grounded in the data and tailored to the research questions. It is designed to ensure a systematic and consistent treatment of all units of data (e.g. transcripts of interviews) across multiple team members and allows for the analytical framework to be refined and modified in the early stages of its use. We would use a mix of deductive and inductive approaches to analysis and coding. Before analysis begins, the framework will be set up with themes drawing upon the implementation dimensions and the IPE research questions, i.e. using a deductive approach. However, during analysis or data collection, evaluators may identify additional issues or questions that could contribute useful insights so these would be incorporated into the framework, i.e. an inductive approach.

Information from each interview will be entered as a separate row in the framework and represents one perspective. Framework allows full within case analysis (looking in detail at each individual row) and between case analysis (comparing individual rows and groups of rows), as well as thematic analysis. Verbatim quotes and evaluator notes are included in the

framework under relevant themes or questions. The context of the information is retained, and the transcript location is noted, so that it is possible to return to a transcript to explore a point in more detail or to extract further text where a verbatim quotation is particularly long. Organising the data in this way allows us to compare the full range of experiences and accounts and patterns across different groups of people.

For quality assurance purposes, one researcher will develop the framework, and a second researcher will check to ensure the appropriate themes are included.

### **Statistical analysis**

Quantitative analysis of the parent survey at both baseline and endline will include basic statistical tests to identify changes in practice or behaviour, using descriptive frequencies and crosstabs, as well as inferential statistics. Data will be analysed to compare intervention and control groups and controlling for change over time (if we have sufficient repeated measures data). This would be achieved using an ANCOVA model or equivalent non-parametric design. Further details will be provided in the SAP.

These findings would then be triangulated with findings from qualitative and monitoring data to explore what factors may contribute to the findings or highlight areas for closer consideration where findings from different sources presenting conflicting pictures.

The quality assurance processes for the quantitative data will include one analyst cleans and codes the data and then runs the analysis. Another analyst will independently check and verify the coding and output.

### **Delivery information**

Information provided by the delivery team about the programme and monitoring its delivery would be used to provide broader insights into delivery that can be used to enrich other analyses.

The delivery team will be asked to provide information about the current price of the intervention for settings/individuals outside of the trial, whether they have made any changes to the intervention during delivery that would mean that the price would be different for the intervention as it has been delivered during the trial, whether they anticipate any change to the price for future delivery, and whether they anticipate any change in the price of the intervention if it were to be delivered on a larger scale. Further detail is provided in the cost evaluation section of the protocol.

Table 5 presents an overview of how the IPE research methods and questions feed into the IPE dimensions identified for this study.

Table 5: IPE methods overview

IPE dimension	RQ addressed	Research methods	Data collection methods	Sample size and sampling criteria	Data analysis methods
Fidelity – the extent to which implementers adhere to intended treatment model	RQ1 RQ7	IDEA Workshop/set-up meeting	Theory of Change development	N.A.	Theory of Change
		Desk review.	Reviewing intervention materials – content of example texts  Knowledge from previous trial	N.A.	N.A.
		<b>Parent endline treatment group survey</b>	Online survey – did parents receive 3 texts per week	Treatment group approx. – 1885 parents (145 schools x 26 pupils per school = 3770 / 2)	Descriptive frequencies, crosstabs
		<b>Monitoring information</b>	Delivery team data on texts sent – did the delivery team send 3 texts per week	3770 parents	Descriptives.
		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
Adaptation – the nature and extent of changes made to the intervention.  Would anything need to be adapted for roll out?	RQ7	<b>Parent endline treatment group survey</b>	Online survey	Treatment group approx – 1885 parents (145 schools x 26 pupils per school = 3770 / 2)	Descriptive frequencies, crosstabs
		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
Dosage – how much of the intended intervention has	RQ1	<b>Parent endline treatment group survey</b>	Online survey	Treatment group approx – 1885 parents (145 schools x 26 pupils per school = 3770 / 2)	Descriptive frequencies, crosstabs

been delivered and/or received		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
		<b>Monitoring information</b>	Delivery team data on texts sent – did the delivery team send 3 texts per week	3770 parents	Descriptives
Quality – How well different component of the intervention are delivered	RQ3	<b>Parent endline treatment group survey</b>	Online survey	Treatment group approx – 1885 parents (145 schools x 26 pupils per school = 3770 / 2)	Descriptive frequencies, crosstabs
		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
Reach – the rate and scope of participation	RQ3	<b>Parent endline treatment group survey</b>	Online survey	Treatment group approx – 1885 parents (145 schools x 26 pupils per school = 3770 / 2)	Descriptive frequencies, crosstabs
		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
		<b>Monitoring information</b>	Delivery team data on texts received	3770 parents	Descriptives
Responsiveness – the degree to which participants engage with the intervention	RQ2 RQ3	<b>Parent endline treatment group survey</b>	Online survey	Treatment group approx – 1885 parents (145 schools x 26 pupils per school = 3770 / 2)	Descriptive frequencies, crosstabs
		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
Programme differentiation – the extent to which intervention activities can be distinguished from other existing practice	RQ3 RQ4	<b>Baseline and endline parent surveys</b>	Online survey – capturing usual practice at baseline and any ongoing similar practice throughout.	3770 parents	Descriptive frequencies, crosstabs
		<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis

Monitoring of control groups – determination of the counterfactual	RQ4	<b>Baseline and endline parent surveys of control group</b>	Online survey –	3770 parents	Descriptive frequencies, crosstabs
Perceived impact – perception from stakeholders on whether the programme achieved its intended outcomes	RQ8	<b>Endline parent surveys in treatment group</b>	Online survey –	1885 parents	Descriptive frequencies, crosstabs
Cost – perceived cost of the programme and whether considered worthwhile	RQ6	<b>School staff survey</b>	Online survey/	145	Descriptives Thematic
Context/moderators – perceived factors that may influence delivery or impact	RQ5	<b>Endline parent surveys in treatment group</b>	Online survey –	1885 parents	Descriptive frequencies, crosstabs
	RQ8	<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis
Mediators – perceived factors that have caused the desired outcomes	RQ5	<b>Endline parent surveys in treatment group</b>	Online survey –	1885 parents	Descriptive frequencies, crosstabs
	RQ8	<b>Parent interviews</b>	In depth exploration with parents	35 parents	Framework analysis, thematic analysis

## Cost evaluation design

Tips by Text is a relatively low-cost intervention. The efficacy trial report estimated that the recurring costs (consisting of message delivery) amounted to £2 per child per year, while the cost of setting up and monitoring the intervention was estimated at £45 per child per year (Stokes et al., 2022). The overall cost of the intervention was therefore estimated at £47 per child per year. There are few hidden costs as school staff are not involved in the intervention, except for keeping and updating parents' phone numbers of parents - an activity which they may have to carry out anyway as part of their work. Similarly, the time demanded of parents was not extremely high. Parents reported spending a few minutes per day in the recommended activities.

To calculate the cost of the intervention we will follow the principles set out in the EEF cost evaluation guidance (EEF, 2023), and will estimate costs from the perspective of the school. We will present cost as per pupil, calculated as if the intervention was implemented over a three-year period. Note that in a scale-up scenario, the intervention could be delivered by a school or by a third-party agency hired for this purpose. We will consider the costs potentially incurred by both options.

Costs will be separated into pre-requisites, start-up costs and recurring costs. From a school perspective there are no pre-requisite or start-up costs. There is an initial cost in developing the programme, which is not incurred by the school and is therefore not included. Any research cost or subsidy provided by EEF is not included among costs either.

The recurring costs will therefore include:

- the cost of setting up the programme in the school: the school will have to collect parents telephone numbers and the messaging firm will set these up in the text messaging system
- the cost of delivering the text messages: this is simply the cost of sending a text message at current rates incurred by the messaging firm
- the cost of monitoring message delivery: the messaging firm will have to send reminders when a text message fails to deliver and will have to update parents' telephone numbers with support from the schools when these change

We will use the ingredients approach for data collection, in which costs of all resources required to implement the programme regardless of who incurs the costs (but excluding research costs) are included. However, we anticipate that nearly all the costs of the intervention will be incurred by the messaging firm as illustrated above.

Costs incurred by the delivery team do not represent a good approximation of the costs that would be incurred by a school, or by another party operating on behalf of multiple schools. The project costs to be incurred by a school wishing to adopt the programme in the future can be obtained in two ways. The first, consists of contacting a firm other than BIT specialised in providing text messages and obtaining a quote of the cost of the programme under different assumptions with regards to its coverage. The scale of operation of the interventions is the most critical one. Since a component of the set-up costs is fixed (such as designing the messaging and setting up the software), the unit cost of each message is likely to decrease with the number of parents reached. In case a messaging firm cannot be identified, a second method consists of using current market prices for the identified ingredients. For example, the evaluation team could obtain estimates of the market cost of a text message, of the market cost of the staff time required to collect and update parents' contact details, of the opportunity costs to parents etc, thus obtaining an overall cost of the intervention.

As much of the cost of the intervention is incurred by the messaging firm, we do not plan to conduct a cost survey in schools. Rather, we will conduct interviews with School finance office staff from a random sample of intervention schools. The interviews will obtain estimates of the school staff time required by the intervention and will reveal any unanticipated cost incurred by the schools.

As for the social cost of the intervention, it is expected that parents will contribute to the intervention through their own time. Interviews conducted for the evaluation of the efficacy trial suggest that parents spent on average 20 minutes per week on the activities recommended by the intervention. The amount of parents' time required appears to be relatively small, amounting to about a total of 10-15 hours per child over the course of the programme. The costs incurred by parents, estimated at the available opportunity cost, can be used to calculate the social cost of the intervention.

Finally, to calculate cost-per-child, we will first estimate the total cost per school if the programme were to be implemented for three consecutive years, and then we will work out the per year cost. To obtain a child-level cost estimate, the total cost will be divided by the number of children per school who would be expected to benefit; we intend to use the actual number of children attending schools in the trial for this purpose. All assumptions used in the cost evaluation will be clearly stated in the report.

## **Ethics and registration**

The proposal has been approved by the NIESR ethics committee after careful consideration and with no issues arising from the committee's deliberations.

The project team prepared an information sheet for schools, which provides an overview of the project, and the first stage is for schools to express interest in the trial by contacting the delivery team directly. Following this, to participate in the trial, schools will be asked to sign a memorandum of understanding, setting out in detail what participation in the project involves and the roles of all parties. Schools will then be asked to distribute information sheets to parents of eligible children, which will clearly explain the evaluation, and how parents can withdraw their child from the evaluation at any time. Parents have a right to withdraw from the trial at any point with no detrimental consequences and data pertaining to that child will not be used. Verbal information will be provided prior to any research encounter, and ethical consent will be confirmed. Verbal information will make clear the purpose of the research, how the data will be used, when the data will be deleted and how to withdraw data. Privacy notices including details of the data being collected will also be provided to schools and parents.

The protocol is registered with the ISRCTN registry, with study registration number ISRCTN48644807: "Evaluation of Tips by Text (a text messaging programme to promote language development in early years)".

## **Data protection**

We recognise the utmost importance of data protection and are committed to complying with the Data Protection Act 2018 and the GDPR. The evaluation will involve the collection of personal data in relation to participating children, as well as personal data relating to parents of participating children in reception year.

A DPIA has been conducted for the evaluation by the evaluation team, which has been reviewed by NIESR's DPO and will be kept under review through the duration of the project. A separate DPIA is also being conducted by the delivery team.

The evaluation team's legal basis for processing personal data is legitimate interests. NIESR considers that it has a legitimate interest in processing the personal data to conduct the Tips by Text evaluation (with undertaking research and evaluation activities one of NIESR's core purposes), the processing identified is necessary in order to meet this purpose, and there are minimal impacts of the processing on the individuals involved (and which we have taken steps to ensure).

The special category data to be processed for this evaluation relates to gender. NIESR's legal basis for processing special categories of personal data is Article 9j) 'archiving, research and statistics.' Processing personal information to home language is necessary for the conduct of the evaluation.

A data sharing agreement has been developed by the evaluation team (including Qa as our assessment partner), the delivery team and EEF, stating which data will be shared by who, how and why.

A privacy notice, drafted in collaboration with the delivery team and EEF, explains how information collected from participants will be used and stored, and communicates to participants their right to withdraw from data processing. All participants interviewed for the research will be asked for their consent to indicate that they understand the aims of the research and agree to the interview being recorded and transcribed, and given assurance of anonymity. Schools will sign an MOU at the start of the project clearly laying out the requirements of participation and how data will be used, shared and stored.

The privacy notice also explains that the personal data described may be shared with third parties. These third parties may include trusted academic partners and subcontractors who may process information on behalf of the evaluation and delivery teams to carry out work such as transcribing interviews and aspects of programme delivery. In all cases, the evaluation and delivery teams will ensure that these third parties enter into appropriate data processing agreements and that they keep the personal data secure and confidential.

The evaluation team and the delivery team will be data controllers of the personal data collected during the evaluation, while our data collection partner will be data processor. At the end of the project (once the evaluation report is finalised), the evaluation team will upload the project data to the EEF archive; this will take place through the ONS Secure Research Service. Once the data are in the archive, the EEF will become the sole data controller, and the Fischer Family Trust (FFT) will be the data processor. For the purposes of research and archiving, the responses will be linked with information about the children from the National Pupil Database (NPD) and shared with the Department for Education, the EEF, the EEF's archive manager, the Office for National Statistics and potentially other research teams. NIESR will delete personal data held in relation to the project six months following publication of the final evaluation report.

## Personnel

**Edoardo Masset** (Associate Research Director, NIESR) is the principal investigator. He will be responsible for the overall evaluation study. He will lead the evaluation team, design the study, lead and carry out the analysis of the data, and write the final report. He will also support the coordination between the evaluation team, the delivery team, and the data collection team. Edoardo is an expert of impact evaluation methods. He has been the technical lead of several experimental and quasi-experimental evaluation and has great experience in managing evaluations of complex interventions. Prior to joining NIESR he was associate professor at the London School of Hygiene and Tropical medicine, where he led a an FCDO-funded programme promoting and supporting the development and use of new evaluation methods in challenging environments.

**Professor Claudine Bowyer-Crane** (NIESR Fellow and University of Sheffield). Claudine is a psychologist with an interest in children's reading and language development, particularly children learning English as an Additional Language. She has extensive with expertise in designing and evaluating interventions to support children's language development. Claudine will advise the team on the identification, use, and analysis of language skills indicators.

**Jasmin Rostron** (Associate Social Researcher, NIESR), with other NIESR colleagues, will be responsible for conducting interviews with practitioners, setting leads and trainer/mentors. She will also contribute to analysis and writing the reports. Jasmin is a qualitative researcher, experienced in interviews, focus groups and evaluations. She has led on two implementation and process evaluations for What Works for Children's Social Care (WWCSC) trials in schools. This involved interviewing designated safeguarding leads, middle and senior leaders. Additionally, she has worked with Refugee Education UK on an ONS-funded qualitative project, on the experiences of displaced young people, accessing education and healthcare.

**Anneka Dawson** (Head of pre-16 education and the IES Trials Unit, on Maternity Leave Sept 2023-November 2024) was Co-PI on the efficacy trial and will provide expert advice to the project team. Anneka has vast experience in education and family research over 15 years. Anneka was co-PI on the efficacy trial and has also directed a number of other EEF projects, in addition to consultancy work for the EEF. She is currently directing a number of early years projects for Sutton Trust and the Mercers Foundation. She was formerly the senior evaluation manager at the EEF with responsibility for Early Years research.

**Susie Bamford** (Principal Research Fellow, IES), is the Project Director which means she will support the IPE team on all aspects of the project, providing input and quality assurance on all research documents, analysis and reporting processes. Susie is a quantitative researcher and chartered psychologist with considerable expertise and experience designing and delivering quantitative, mixed methods research, and Randomised Controlled Trials (RCTs) spanning 20 years. Her experience spans several disciplines including psychology, experimental medicine, and education. This includes several years as a postdoc in academia running internationally recognised trials and a position at the National Foundation for Educational Research (NFER) specialising in RCTs and education and employment. Here she led several EEF trials. She then spent 4 years as an expert independent consultant

advising clients on RCTs and in October 2022 she joined the Institute for Employment Studies as a Senior Research Fellow.

**Jade Talbot** (Research Officer, IES) will contribute towards all aspects of the quantitative and qualitative data collection and analysis (surveys, interviews, analysis and writing up). Jade has experience of applied social research in education settings, including engaging with multiple stakeholders such as children, parents, and teachers. Jade is currently working on a number of projects focused on the development of language and literacy skills. In the EYFS, this includes the EEF Tales Toolkit evaluation and the Mercers-Peeple STEM programme evaluation. At secondary stage, Jade is working on the DfE MFL Programmes evaluation, as well as the Apps for Good evaluation.

## Risks

The main risks facing the intervention are discussed in the table below.

Issue/risk	Action to address issue/reduce risk	Likelihood	Impact
Contamination of the control group	Participating parents will be informed of the experimental nature of the intervention and asked not to share text messages with non-participating parents	Medium	High
Poor participation in baseline assessment and high attrition.	Requirements of the study, including assessments, will be made clear to schools and to parents at recruitment.  Ensuring flexibility in the scheduling of assessments.  Understanding patterns of attendance	Medium	High
Large drop-out rates in schools with poor capacity.	The data collection team will conduct mopping-up visits in order to administer the largest possible number of assessments at both baseline and follow-up. Dates of assessments will be shared with parents and practitioners in advance to the extent it will be possible given the tight data collection timeline.	Medium	High
Reluctance of parents and school staff to participate in interviews	Reminders will be developed to increase parents' participation in online surveys.	Low-Medium	High
Data collection process does not adhere to agreed protocol	The evaluation team will meet periodically with the data collection team to define a feasible and detailed data collection plan. A member of the evaluation team will make spot visits to schools where data collection takes place to check adherence to the protocol. This will occur in the first week of data collection in order to be able to make any remedial correction to the protocol. The evaluation and the data collection team will	Low-medium	High

	be constantly in contact during critical phases of data collection.		
Reluctance of schools to participate	The delivery team will be responsible for recruitment and will start recruitment early.	Low-medium	Medium
NPD data cannot be accessed	DfE will be contacted early for access. Evaluation team will obtain access and training to use the data.	Low-medium	Low

## Timeline

<b>Dates</b>	<b>Activity</b>	<b>Staff responsible/ leading</b>
Mar-Jul 2024	Recruitment of schools	Delivery team
Jun 2024	Training of QA on assessment tools	Evaluation team
Jun 2024	Piloting of assessment tools	Evaluation team
Sep 2024	Training of assessors on assessment tools	Data collection team
Sep 2024	1 <sup>st</sup> , 2 <sup>nd</sup> and 3 <sup>rd</sup> week (wc Sept 2 <sup>nd</sup> ,9 <sup>th</sup> and 16 <sup>th</sup> ) Collection of pupil data from schools, parent withdrawal (1 weeks from sending the letters)	Delivery team
Sep-Nov 2024	From 3 <sup>rd</sup> week (wc Sept 16 <sup>th</sup> ), Baseline assessment	Data collection team
Oct 2024	From 5 <sup>th</sup> week (wc Sept 30 <sup>th</sup> ), Baseline survey of parents and school staff	Evaluation team
Oct 2024	Randomisation (during half term)	Evaluation team
Nov-Jul 202	Programme commences on 11 <sup>th</sup> November and runs for 9 months	Delivery team
Feb-Mar 2025	Parents IPE interviews 1 <sup>st</sup> round longitudinal cohort	Evaluation team
Jun-Jul 2025	Post-assessment data collection	Data collection team
Jun-Jul 2025	Endline survey of parents and teachers, and parent IPE interview 2 <sup>nd</sup> round longitudinal and endline cohort only	Evaluation team
Jul 2025	Programme ends	Delivery team
Aug-Dec 2025	Data analysis and draft report	Evaluation team
Nov 2025	Linkage to NPD data and further analysis	Evaluation team
Feb 2026	First draft of evaluation report	Evaluation team
May 2026	Final evaluation report	Evaluation team

## Appendix 1: Tips by Text Data collection plan

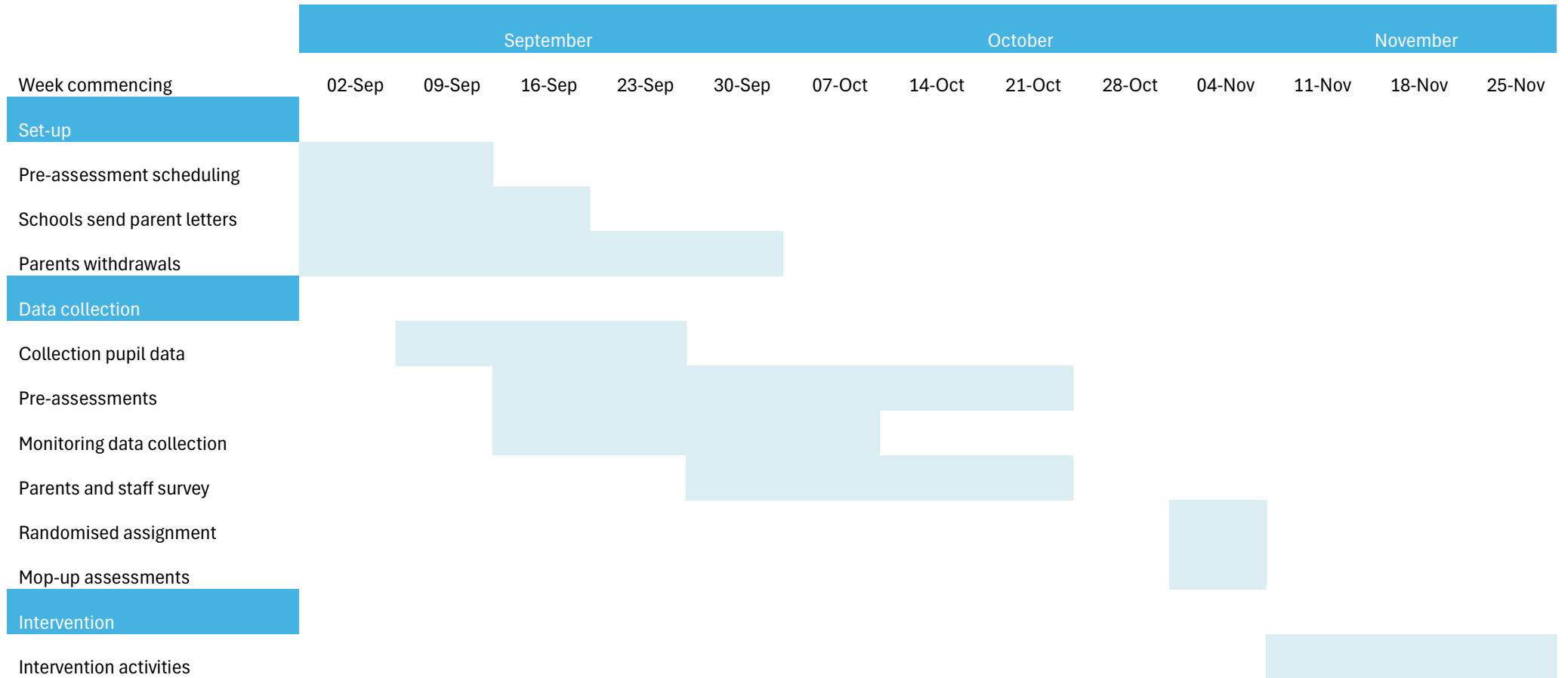
Dates	Activity	Responsibility
Feb - July 2024	Recruitment of schools	Delivery team
June - July 2024	Provide information on recruited schools (including school name, postcode, contact details, %FSM,)	<p>Delivery team share with NIESR who share with IES and Qa (assessment team) for planning purposes</p> <p>BIT plan to send 3 emails in Summer 2 to cover a. making sure schools have GDPR sorted b. to ask schools if project lead is changing and c. what schools can expect in sept/oct – outlines key dates.</p>
June - July 2024	Contact schools in preparation for assessment period	<p>NIESR follow up with schools</p> <ul style="list-style-type: none"> <li>• provide schools with information on evaluation activities and timeline</li> <li>• instructions for data collection and how to set up for using sharepoint (for data sharing)</li> <li>• (in multi-form entry schools) ask for information of classes (for random selection of 1 class for testing)</li> <li>• Provide email and phone number for schools to contact if any issues</li> </ul> <p>Assessment team will also contact schools in preparation for Sept – start booking in provisional assessment dates</p>
June 2024	<p>Training in Language Screen (primary outcome measure)</p> <p>Training in Heads, Knees, Shoulders and Toes revised (HKTS-R) (secondary outcome measure)</p>	Language Screen OxEd (Language Screen developer) run session for Senior Qa staff, NIESR and IES (5 <sup>th</sup> June 2024 online) on the use of language screen, its validity, and its software
June - July 2024	Draft baseline surveys of parents and staff*	<p>IES lead with input from NIESR and EEF.</p> <p>First draft shared with NIESR 14/6/24 for review.</p> <p>IES will share first draft with EEF on June 28<sup>th</sup>.</p> <p>EEF send comment to IES wc July 8<sup>th</sup> for revisions tbc</p> <p>IES make revisions final two weeks of July</p>
Jul 2024	School contacted to start arrangements for scheduling assessments.	Qa contacts the schools by email or by any other contact detail available in an attempt to schedule assessments in advance.
Aug 2024	1 class is randomly selected in multi-class settings and If there are more than 30 children in a class, 30 children are randomly selected as soon as class size is known.	NIESR will contact schools over July-August period to explain one class only will be selected and a maximum of 30 children per class will be selected, and that both will be selected randomly. The selection will be made in August and

		communicated to the schools as soon as it is made.
September 2024	Sharing info sheet and withdrawal forms with schools	Delivery team shares forms with schools
Sep 2024 (Week 1-2)	In person training of assessors in the 4 regions. in Language Screen (primary outcome measure) Training in Heads, Knees, Shoulders and Toes revised (HKTS-R) (secondary outcome measure)	Senior Qa staff will train the assessors on how to use the test. LanguageScreen requires minimal training and Qa have used HKTS-R assessment in the past and are able to train the assessors. Wednesday 4th September – Southampton Thursday 5th September – London Likely dates are Wed 11th September – York Wed 11th September – Manchester Thuy 12th September – Liverpool
Sep 2024 (Week 1-4)	Schools send parents letters out Parent withdrawal period starts (Allow 1 week from when sent letter). Friday of Week 2 is the last date for sending letters (Sept 13 <sup>th</sup> ), Friday of Week 3 is last day for withdrawal (Sept 20 <sup>th</sup> ).	Schools with support from the delivery team who confirm with schools that letters have gone out and what date.  BIT share this with NIESR and assessment team
Sep 2024 (Week 1)	Co-ordinating impact data collection in schools – BIT will share with schools an excel template for pupil data collection (including UPN, names, gender, dob, FSM, parents contact details (mobile telephone number and email addresses) in the first week of September – NB in the template schools report the number of children withdrawn and only children not withdrawn are included).	BIT
Sep 2024 (Week 1 - 5) meaning until Friday 4 <sup>th</sup> of October	Monitoring pupil data collection in schools.	BIT follows up with schools - where pupil data template has not been returned or if there are errors/blanks in the data template – daily updates are provided and the teams are allocated specific schools/areas to telephone/email key contacts re template remind leads to check spam, complete template, raise any issues.
Sep 2024 (Week 1 -Week 5)	Sharing pupil data with Qa to plan data collection.	BIT shares pupil data with Qa and NIESR. NIESR share parent data with IES (name, email address and mobile phone number)
Sep 2024 (Week 1 - Week 2)	School contacted to schedule pre-assessments (info exchange to plan pre-assessments)	Qa contacts the schools by email or by any other contact detail available. In case of no response from the school within 2 working days, they will make a second attempt. Failing the second attempt after 2 working days, Qa will notify NIESR who will be in touch with the school directly.

Sep 2024 (Week 5) – Nov (Week 1)	Surveys of parents and staff sent out	IES emails invite includes hyperlink to survey for completion by parents and staff. For parents <b>only</b> - if email address not provided by school, SMS will be sent with information and link to survey.
Sep 2024 (Week 4) – Oct 2024 (Week 3)	Pre-assessment data collection of an average of 26 tests in 145 settings (total 3770 tests). Up to 30 tests will be conducted in each school. 60 assessors are deployed (15 in each region). Under ideal conditions pre-assessment takes 25m to complete. One assessor can cover 1 school in about 3 days (assume 5hrs in school and about 25m per assessment), then 15 assessors can cover 36 schools in about 7 working days. However, to account for various delays, cancellations and revisits we are allowing a data collection period of 6 weeks corresponding to 30 working days (this includes the 1 mopping up week in November below)	Qa will test all children in the selected classes, delivering an expected average of 26 pupils per class. Qa will test children from the pupil data list collected by NIESR.  There is an estimated 10% rescheduling of visits due to school cancellations.  LanguageScreen assessment data are automatically uploaded by the application as the test is completed, they are securely stored by the application, and accessible to the evaluation team only. NIESR will monitor progress in data collection, completeness and accuracy.  The HKTS-R data are collected on paper, transcribed on a spreadsheet and manually uploaded by Qa, and they will be shared with NIESR using Sharepoint. Qa have systems in place to monitor completeness and accuracy of collected data.
Nov (Week 1)	Mop-ups: school are revisited to complete target number of assessments	Qa re-visits schools. Estimated mop-up of 10%.
Nov (Week 1)	Randomised assignment of pupils to the intervention	NIESR shares outcome with Delivery team
Nov 2024 (Week 2 - July 2025)	9-month intervention period	Delivery team
Feb- March 2025	Parent interviews: first round (longitudinal cohort)	IES lead with support from NIESR (parents opt-in at baseline survey) Sample will be split evenly, with NIESR and IES each having half to sample from. Each team will carry out half the interviews via telephone/online (as preferred by parent) Payment of £40 following each interview (online shopping voucher)
April 2025	Draft endline surveys of parents and staff*	IES lead with input from NIESR and EEF.
May 2025	Programme and test endline surveys of parents and staff*	IES with input from NIESR and EEF using Snap survey programme
June -July 2025	Post-assessments after May half term – June 2 <sup>nd</sup> to July 25 <sup>th</sup>	Qa

June -July 2025	Parent interviews: second round (longitudinal cohort)	IES lead with support from NIESR – each team will follow up on the parents they previously interviewed. NIESR and IES do half each. In case of attrition, where parents do not want be interviewed again, more parents who opted in at the baseline survey will be asked to be interviewed Via telephone/online (as preferred by parent. Payment of £40 following each interview (online shopping voucher)
June -July 2025	Parent interviews: endline cohort only	IES lead with support from NIESR – each team will follow up on parents who opted into interview in baseline survey (not including those in longitudinal cohort or have indicated they do not want to take part). NIESR and IES do half each. Via telephone/online (as preferred by parent. Payment of £40 following each interview (online shopping voucher)
June -July 2025	Follow-up interview with parents dropping out	IES only
June – July 2025	Endline surveys of parents and school staff	IES
August - December 2025	Analysis and draft reporting	IES and NIESR
Nov 2025	Linkage to NPD data (including EYFSP)	NIESR

## Appendix 2 Gantt chart of baseline data collection



### Appendix 3 Tips by Text communication plan

Activity	Responsible	Shared with	Deadline
Finalise and share MoU and school information sheet.	BIT	Evaluation team	completed
Finalise and share parent information and withdrawal form, privacy notice, staff privacy notice	Evaluation team	BIT	completed
Finalise and share school data sharing agreement and DPIA	NIESR	BIT, IES, Qa	completed
Recruitment of schools	BIT	Schools	completed
MoU and data sharing acknowledgement signed	Schools	BIT	12 <sup>th</sup> Feb to 26 <sup>th</sup> July 24
Provide information on recruited schools (including multi-class schools)	BIT	NIESR	19 <sup>th</sup> July 24
Provide information on recruited schools for planning purposes (including selected class in case of multi-class schools)	NIESR	Qa	26 <sup>th</sup> July
Share with schools template for collection of pupil data (names, dob, FSM, parents contacts)	BIT	Schools	2 <sup>nd</sup> Sep 24
Share with schools parent information and withdrawal form	BIT	Schools	2 <sup>nd</sup> Sept24
Parents information forms are sent-out (withdrawal option only is returned) by the 6 <sup>th</sup> of September, but schools will be allowed to send the letters until the 20 <sup>th</sup> of September	Schools	Parents	20 <sup>th</sup> Sep 24
Return template for collection pupil data excluding withdrawn children	Schools	BIT	27 <sup>th</sup> Sep
Share parents' details for intervention and surveys activities	BIT, NIESR	NIESR IES	30 <sup>th</sup> Sep 24
Periodically contact schools regarding pupil data and withdrawals	BIT	Schools	1 <sup>st</sup> Nov 24
Periodic updates on parents and schools withdrawals	BIT	NIESR	1 <sup>st</sup> Nov 24
Daily reports on data collection activities	BIT	NIESR, QA, IES, EEF	1 <sup>st</sup> Nov 24

--	--	--

Conduct baseline assessments and share data	Qa	NIESR	1st Nov 24
Conduct parent and school staff baseline surveys	IES	NIESR	8 <sup>th</sup> Nov 24
Randomly assign children to the intervention	NIESR	BIT	8 <sup>th</sup> Nov 24
Deliver the intervention (text messages)	BIT	Parents	11th Nov - 25th Jul 25
Conduct IPE parent interviews (longitudinal cohort 1 <sup>st</sup> interview)	IES	NIESR	28 <sup>th</sup> Mar 25
Periodic updates on parents and schools withdrawals	BIT	NIESR	25th Jul 25
Keep records of parents who have withdrawn	BIT	IES	25 <sup>th</sup> Jul 25
Keep records of messages delivered and undelivered	BIT	NIESR	25 <sup>th</sup> Jul 25
Conduct endline assessments	Qa	NIESR	25th Jul 25
Share endline assessment data	Qa	NIESR	8 <sup>th</sup> Aug 25
Conduct parent and school staff endline survey	IES	NIESR	25 <sup>th</sup> Jul 25
Conduct IPE parent interviews (longit cohort follow-up and endline cohort only)	IES	NIESR	25 <sup>th</sup> Jul 25

## References

- Attanasio, O., Cattan, S. and Meghir, C. (2022) 'Early Childhood Development, Human Capital, and Poverty', *Annual Review of Economics*. Available at: <https://doi.org/10.1146/annurev-economics-092821-053234>.
- August, D. and Shanahan, T. (2006) *Developing literacy in second language learners: Report of the national literacy panel on language minority children and youth*. Edited by D. August and T. Shanahan. Mahwah, NJ: Erlbaum.
- Bishop, D.V.M., Nation, K. and Patterson, K. (2014) 'When words fail us: Insights into language processing from developmental and acquired disorders', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Available at: <https://doi.org/10.1098/rstb.2012.0403>.
- Borenstein, M. *et al.* (2009) *Introduction to Meta-Analysis, Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons, Ltd. Available at: <https://doi.org/10.1002/9780470743386>.
- Cabell, S.Q. *et al.* (2019) 'Impact of a Parent Text Messaging Program on Pre-Kindergarteners' Literacy Development', *AERA Open*, 5(1). Available at: <https://doi.org/10.1177/2332858419833339>.
- Cattan, S. *et al.* (2022) 'Early childhood inequalities', *IFS Deaton Review of Inequalities* [Preprint].
- Cortes, K.E. *et al.* (2021) 'Too Little or Too Much? Actionable Advice in an Early-Childhood Text Messaging Experiment', *Education Finance and Policy*, 16(2). Available at: [https://doi.org/10.1162/edfp\\_a\\_00304](https://doi.org/10.1162/edfp_a_00304).
- Damgaard, M.T. and Nielsen, H.S. (2018) 'Nudging in education', *Economics of Education Review*. Available at: <https://doi.org/10.1016/j.econedurev.2018.03.008>.
- Doss, C. *et al.* (2019) 'More Than Just a Nudge', *Journal of Human Resources*, 54(3). Available at: <https://doi.org/10.3368/jhr.54.3.0317-8637r>.
- Doss, C. *et al.* (2022) 'Engaging girls in math: The unequal effects of text messaging to help parents support early math development', *Economics of Education Review*, 88. Available at: <https://doi.org/10.1016/j.econedurev.2022.102262>.
- EEF (2022) *Statistical Analysis Guidance for EEF Evaluations*. London: EEF.
- EEF (2023) *Cost evaluation guidance for EEF evaluations*. London: EEF.
- Gonzales, C.R. *et al.* (2021) 'The Head-Toes-Knees-Shoulders Revised (HTKS-R): Development and psychometric properties of a revision to reduce floor effects', *Early Childhood Research Quarterly*, 56. Available at: <https://doi.org/10.1016/j.ecresq.2021.03.008>.

Gooch, D. *et al.* (2016) 'The development of executive function and language skills in the early school years', *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 57(2). Available at: <https://doi.org/10.1111/jcpp.12458>.

Haushofer, J. and Fehr, E. (2014) 'On the psychology of poverty', *Science*. Available at: <https://doi.org/10.1126/science.1232491>.

Heckman, J.J. and Masterov, D. V. (2007) 'The productivity argument for investing in young children', in *Review of Agricultural Economics*. Available at: <https://doi.org/10.1111/j.1467-9353.2007.00359.x>.

Hulme, C. *et al.* (2023) . 'LanguageScreen: The development, validation and standardization of an automated language assessment ', *Available online: https://osf.io/preprints/osf/gq6fh* [Preprint].

Hurwitz, L.B. *et al.* (2015) 'Supporting Head Start parents: impact of a text message intervention on parent–child activity engagement', *Early Child Development and Care*, 185(9). Available at: <https://doi.org/10.1080/03004430.2014.996217>.

Imbens, G.W. and Rubin, D.B. (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences, Causal Inference: For Statistics, Social, and Biomedical Sciences an Introduction*. New York: Cambridge University Press. Available at: <https://doi.org/10.1017/CBO9781139025751>.

Jeynes, W. (2012) 'A Meta-Analysis of the Efficacy of Different Types of Parental Involvement Programs for Urban Students', *Urban Education*, 47(4). Available at: <https://doi.org/10.1177/0042085912445643>.

Kaushanskaya, M. *et al.* (2017) 'The relationship between executive functions and language abilities in children: A latent variables approach', *Journal of Speech, Language, and Hearing Research*, 60(4). Available at: [https://doi.org/10.1044/2016\\_JSLHR-L-15-0310](https://doi.org/10.1044/2016_JSLHR-L-15-0310).

Kieffer, M.J. (2008) 'Catching Up or Falling Behind? Initial English Proficiency, Concentrated Poverty, and the Reading Growth of Language Minority Learners in the United States', *Journal of Educational Psychology*, 100(4). Available at: <https://doi.org/10.1037/0022-0663.100.4.851>.

Mayer, S.E. *et al.* (2020) 'Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together Intervention', *Journal of Human Resources*, 54(4). Available at: <https://doi.org/10.3368/jhr.54.4.0617.8835R>.

McClelland, M.M. *et al.* (2021) 'The Head-Toes-Knees-Shoulders Revised: Links to Academic Outcomes and Measures of EF in Young Children', *Frontiers in Psychology*, 12. Available at: <https://doi.org/10.3389/fpsyg.2021.721846>.

Mol, S.E. *et al.* (2008) 'Added value of dialogic parent-child book readings: A meta-analysis', *Early Education and Development*, 19(1). Available at: <https://doi.org/10.1080/10409280701838603>.

Reese, E., Sparks, A. and Leyva, D. (2010) 'A review of parent interventions for preschool children's language and emergent literacy', *Journal of Early Childhood Literacy*, 10(1). Available at: <https://doi.org/10.1177/1468798409356987>.

Shokrkon, A. and Nicoladis, E. (2022) 'The Directionality of the Relationship Between Executive Functions and Language Skills: A Literature Review', *Frontiers in Psychology*. Available at: <https://doi.org/10.3389/fpsyg.2022.848696>.

Singh, A. *et al.* (2023) *Improving power calculations in educational trials*. Available at: <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/eef-evaluation-reports-and-research-papers/methodological-research-and-innovations/improving-power-calculations-in-educational-trials> (Accessed: 31 January 2024).

Smythe-Leistico, K. and Page, L.C. (2018) 'Connect-Text: Leveraging Text-Message Communication to Mitigate Chronic Absenteeism and Improve Parental Engagement in the Earliest Years of Schooling', *Journal of Education for Students Placed at Risk*, 23(1–2). Available at: <https://doi.org/10.1080/10824669.2018.1434658>.

Stokes, L. *et al.* (2022) *Tips by Text: Evaluation Report*. London: EEF.

Sussman, J.B. and Hayward, R.A. (2010) 'Using instrumental variables to adjust for treatment contamination in randomised controlled trials', *BMJ (Online)*, 340(7757). Available at: <https://doi.org/10.1136/bmj.c2073>.

York, B., Loeb, S. and Doss, C. (2019) 'One Step at a Time: The Effects of an Early Literacy Text-Messaging Program for Parents of Preschoolers', *Journal of Human Resources*, 54(3). Available at: <https://doi.org/10.3368/jhr.54.3.0517-8756R>.